GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 10:15:55 ; Search time 49.36 Seconds
(without alignments)
7254.087 Million cell updates/sec

Title: US-09-603-124B-1
Perfect score: 1581
Sequence: 1 gcaggtaacgcctccacggt.....gacgtaagtgaacaaggcag 1581

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	18	1.1	2289	2	US-08-463-081B-30	Sequence 30, Appl
c 2	18	1.1	2289	2	US-08-461-379A-30	Sequence 30, Appl
c 3	18	1.1	2289	2	US-08-462-390B-30	Sequence 30, Appl
c 4	18	1.1	2289	3	US-08-463-074B-30	Sequence 30, Appl
c 5	18	1.1	2289	3	US-08-465-585C-30	Sequence 30, Appl
c 6	18	1.1	2289	3	US-08-652-446-30	Sequence 30, Appl
c 7	18	1.1	2946	1	US-08-473-122-1	Sequence 1, Appli
c 8	18	1.1	2946	2	US-08-472-478-1	Sequence 1, Appli
c 9	18	1.1	2946	2	US-08-463-081B-7	Sequence 7, Appli
c 10	18	1.1	2946	2	US-08-461-379A-7	Sequence 7, Appli
c 11	18	1.1	2946	2	US-08-462-390B-7	Sequence 7, Appli
c 12	18	1.1	2946	3	US-08-463-074B-7	Sequence 7, Appli
c 13	18	1.1	2946	3	US-08-465-585C-7	Sequence 7, Appli
c 14	18	1.1	2946	3	US-08-652-446-7	Sequence 7, Appli
c 15	17	1.1	300	2	US-08-308-494A-16	Sequence 16, Appl
c 16	17	1.1	309	1	US-08-467-393-3	Sequence 3, Appli
c 17	17	1.1	322	3	US-08-434-000A-11	Sequence 11, Appl
c 18	17	1.1	322	4	US-09-312-157-11	Sequence 11, Appl
c 19	17	1.1	324	3	US-08-483-749A-7	Sequence 7, Appli
c 20	17	1.1	327	2	US-08-002-324-3	Sequence 3, Appli
c 21	17	1.1	327	5	PCT-US94-00261-3	Sequence 3, Appli
c 22	17	1.1	331	3	US-08-836-561-32	Sequence 32, Appl
c 23	17	1.1	357	4	US-09-171-945-8	Sequence 8, Appli
c 24	17	1.1	360	1	US-08-447-422-13	Sequence 13, Appl
c 25	17	1.1	384	2	US-08-656-586-1	Sequence 1, Appli
c 26	17	1.1	384	4	US-08-619-491-1	Sequence 1, Appli
c 27	17	1.1	384	5	PCT-US95-07302-1	Sequence 1, Appli
c 28	17	1.1	393	2	US-08-116-778E-39	Sequence 39, Appl
c 29	17	1.1	393	2	US-08-438-562-39	Sequence 39, Appl
c 30	17	1.1	393	2	US-08-483-528B-2	Sequence 2, Appli
c 31	17	1.1	393	3	US-08-673-799C-2	Sequence 2, Appli
c 32	17	1.1	462	1	US-07-946-421-25	Sequence 25, Appl
c 33	17	1.1	646	2	US-08-737-129A-3	Sequence 3, Appli
c 34	17	1.1	648	6	5455030-4	Patent No. 5455030
c 35	17	1.1	669	2	US-08-190-199A-66	Sequence 66, Appl
c 36	17	1.1	672	2	US-08-190-199A-62	Sequence 62, Appl
c 37	17	1.1	705	4	US-09-171-945-16	Sequence 16, Appl
c 38	17	1.1	708	2	US-08-190-199A-60	Sequence 60, Appl
c 39	17	1.1	711	2	US-08-190-199A-64	Sequence 64, Appl
c 40	17	1.1	719	3	US-08-279-772A-7	Sequence 7, Appli
c 41	17	1.1	720	1	US-08-061-092A-2	Sequence 2, Appli
c 42	17	1.1	720	3	US-08-902-486-10	Sequence 10, Appl
c 43	17	1.1	720	6	5455030-14	Patent No. 5455030
c 44	17	1.1	729	6	5455030-16	Patent No. 5455030
c 45	17	1.1	732	2	US-08-553-497A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-463-081B-30/c
; Sequence 30, Application US/08463081B
; Patent No. 5871960
; Patent No. 5837487
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression Ther
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; STREET: 444 South Flower St. - Suite 1900
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,081B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,736
; FILING DATE: 10-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-463-081B-30

Query Match 1.1%; Score 18; DB 2; Length 2289;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1333 tgccgtggtgcttgagat 1350
||||||||||||||||||
Db 2113 TGCCGTGGTGCTTGAGAT 2096

RESULT 2
US-08-461-379A-30/c
; Sequence 30, Application US/08461379A
; Patent No. [redacted]
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
; TITLE OF INVENTION: Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; ADDRESSEE: . (B) STREET:One Westlakes-Berwyn
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,379A
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/330,108; 08/104,736
; APPLICATION NUMBER: & 07/796,066
; FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: DART-070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)470-0700
; TELEFAX: (610)470-0701
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-461-379A-30

Query Match 1.1%; Score 18; DB 2; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1333 tgccgtggtgcttgagat 1350
||||||||||||||||||
Db 2113 TGCCGTGGTGCTTGAGAT 2096

RESULT 3
US-08-462-390B-30/c
; Sequence 30, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
; APPLICANT: Smith, K.. A., & Beadling, C.
; TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
; TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; ADDRESSEE: . (B) STREET:One Westlakes-Berwyn
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM; PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,390B
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/330,108
; FILING DATE: 27-OCT-1994
; APPLICATION NUMBER: USSN 08/104,736
; FILING DATE: 10-AUG-1993
; APPLICATION NUMBER: USSN 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: DART-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)407-0700
; TELEFAX: (610)407-0701
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-462-390B-30

Query Match 1.1%; Score 18; DB 2; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1333 tgccgtggtgcttgagat 1350
||||||||||||||||||
Db 2113 TGCCGTGGTGCTTGAGAT 2096

RESULT 4
US-08-463-074B-30/c
; Sequence 30, Application US/08463074B
; Patent No. 6020155
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector an
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; ADDRESSEE: . (B) STREET: 444 South Flower St. - Suite 1
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,074B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 07:44:50 ; Search time 1618.92 Seconds
(without alignments)
16110.766 Million cell updates/sec

Title: US-09-603-124B-1
Perfect score: 1581
Sequence: 1 gcaggtaacgcctccacggt.....gacgtaagtgaacaaggcag 1581

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	1581	100.0	349980	6	AX127150	AX127150 Sequence
	2	1558	98.5	2291	1	AB015023	AB015023 Corynebac
	3	1545.6	97.8	4622	1	BLA242646	AJ242646 Brevibact
	4	1458	92.2	1458	6	AX122446	AX122446 Sequence
	5	1445.2	91.4	1458	6	E28467	E28467 Novel murC
c	6	392.2	24.8	20400	1	AE007068	AE007068 Mycobacte
c	7	392.2	24.8	37586	1	MTCY270	Z95388 Mycobacteri
c	8	392.2	24.8	37586	6	AX191745	AX191745 Sequence
	9	329.2	20.8	348450	1	MLEPRTN4	AL583920 Mycobacte
	10	308.4	19.5	38793	1	MSGB27CS	L78817 Mycobacteri
	11	289.8	18.3	4116	1	AB003132	AB003132 Corynebac
	12	280.8	17.8	5546	1	BLFTSZ	Y08964 Brevibacter
	13	231.8	14.7	24800	1	SC9B1	AL049727 Streptomy
c	14	226.2	14.3	38859	1	MLCB268	AL022602 Mycobacte
c	15	163.6	10.3	329709	1	AP002997	AP002997 Mesorhizo
c	16	162	10.2	294800	1	SME591789	AL591789 Sinorhizo
c	17	141.4	8.9	14545	1	AE008126	AE008126 Agrobacte
c	18	133.8	8.5	195767	1	NMA7Z2491	AL162758 Neisseria
c	19	132.8	8.4	110000	2	LMFLCHR36_27	Continuation (28 o
	20	132.4	8.4	10629	1	AE002398	AE002398 Neisseria
	21	132.4	8.4	349980	6	AX044029	AX044029 Sequence
	22	130.8	8.3	20062	1	AE003919	AE003919 Xylella f
	23	130	8.2	1476	6	AX189061	AX189061 Sequence
	24	130	8.2	1998	1	ECU67892	U67892 Escherichia
	25	130	8.2	2793	1	ECMURGC	X52644 Escherichia
	26	130	8.2	21757	1	AE000118	AE000118 Escherich
	27	130	8.2	28277	1	EC2MIN	X55034 E. coli 2 m
	28	130	8.2	28277	6	AX191720	AX191720 Sequence
	29	130	8.2	111401	1	ECO110K	D10483 E.coli K12
c	30	129.4	8.2	10487	1	AE005922	AE005922 Caulobact
	31	128.4	8.1	281530	1	AP002550	AP002550 Escherich
	32	128	8.1	12518	1	AE005186	AE005186 Escherich
	33	126.6	8.0	1550	1	AF306766	AF306766 Pseudomon
	34	126.6	8.0	1700	1	AF110740	AF110740 Pseudomon
c	35	126.6	8.0	24000	1	AE004856	AE004856 Pseudomon
c	36	119.4	7.6	22201	1	AE004310	AE004310 Vibrio ch
c	37	117.4	7.4	38510	1	AF179611	AF179611 Zymomonas
	38	112.8	7.1	12085	1	U32794	U32794 Haemophilus
	39	106.6	6.7	11584	1	AE006049	AE006049 Pasteurel
c	40	104.2	6.6	15862	1	AE000736	AE000736 Aquifex a
	41	96.2	6.1	19619	6	AX067439	AX067439 Sequence
	42	94	5.9	12394	1	AE001920	AE001920 Deinococc
	43	92.4	5.8	312430	1	RPXX02	AJ235271 Rickettsi
	44	84.6	5.4	1713	1	ECMURG	X52540 Escherichia
	45	84.4	5.3	21000	1	AB052554	AB052554 Shewanell

ALIGNMENTS

RESULT 1
AX127150/c
LOCUS AX127150 349980 bp DNA PAT 11-MAY-2001
DEFINITION Sequence 7066 from Patent EP1108790.
ACCESSION AX127150 AX114121
VERSION AX127150.1 GI:14041138
KEYWORDS .
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Corynebacteriaceae;
Corynebacterium.
REFERENCE 1 (bases 1 to 349980)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7066 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES Location/Qualifiers
source 1. .349980

```
                /organism="Corynebacterium glutamicum"
                /db_xref="taxon:1718"
                /note="Seq 1 to long (3.309.400) split in 11, seq 7066
                2.100.001 2.449.980"
BASE COUNT     80724 a  98367 c  90490 g  80399 t
ORIGIN
```

Query Match 100.0%; Score 1581; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 gcaggtaacgcctccacggtgattgcagacatgattgctgcaactatcaatagccaacac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182723 GCAGGTAACGCCTCCACGGTGATTGCAGACATGATTGCTGCAACTATCAATAGCCAACAC 182664

Qy      61 aactaaaacgaccagctcaacgcaaaggaatagtttaaaggtgaccactccacacttgga 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182663 AACTAAAACGACCAGCTCAACGCAAAGGAATAGTTTAAAGGTGACCACTCCACACTTGGA 182604

Qy     121 ttctgcacaagatattgatctgtcccgcgtccacctcatcggtattggcggagccggaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182603 TTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATCGGTATTGGCGGAGCCGGAAT 182544

Qy     181 gtctggcgttgcccgaatcctgcttgcccgcggtaagacagtcactggttccgatgccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182543 GTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACAGTCACTGGTTCCGATGCCAA 182484

Qy     241 agattcccgcaccttgcttccactccgcgccgtgggagccaccatcgcagtgggacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182483 AGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCCACCATCGCAGTGGGACACGC 182424

Qy     301 tgcggaaaaccttgagctttccggcgaacttcccaccgtcgtggtgacctcttttgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182423 TGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTCGTGGTGACCTCTTTTGCCGC 182364

Qy     361 cattccgcaagacaacccggaacttgttcgtgcacgtgaagaaggcattccggttattcg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182363 CATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAAGAAGGCATTCCGGTTATTCG 182304

Qy     421 tcgctccgatctgttgggcgaattgctggaaggctccacccaggtcttgatcgcgggtac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182303 TCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACCCAGGTCTTGATCGCGGGTAC 182244

Qy     481 ccacggtaagacctccaccacctctatgtctgtggtagctatgcaggcagcgggcatgga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182243 CCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCTATGCAGGCAGCGGGCATGGA 182184

Qy     541 tccaagctttgctatcggcggacagctcaacaaggctggcaccaatgcgcaccatggaac 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182183 TCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGCACCAATGCGCACCATGGAAC 182124

Qy     601 tggtgaggtctttatcgctgaagcagatgaatctgacgcatcgctgctgcgctacaagcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182123 TGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCATCGCTGCTGCGCTACAAGCC 182064

Qy     661 aaatgttgcagtggtcaccaatgtggaaccagaccacctggacttctttaaaacccctga 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182063 AAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTGGACTTCTTTAAAACCCCTGA 182004

Qy     721 agcctacttccaagtgttcgacgatttcgcaggacgcatcaccccgaacggcaagctggt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182003 AGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATCACCCCGAACGGCAAGCTGGT 181944

Qy     781 tgtgtgcctgaacgatcctcacgcagcggagctgggggagaggtctgtccgcaagggtat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181943 TGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAGAGGTCTGTCCGCAAGGGTAT 181884

Qy     841 caagactgttggttacggtaccgctgacgccgtacaggcacaccctgaggttccagcgat 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181883 CAAGACTGTTGGTTACGGTACCGCTGACGCCGTACAGGCACACCCTGAGGTTCCAGCGAT 181824

Qy     901 ggctaccatcgtggattcccaagttgtcgcagaaggcacccgcgccaccatcaacatcga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181823 GGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCATCAACATCGA 181764

Qy     961 tggacaggaagtatctgtgattcttcaaatccctggtgatcacatggtactcaacggtgc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181763 TGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACTCAACGGTGC 181704

Qy    1021 agccgccctgctggccggatacctggtgggtggggacgtcgacaagcttgttgaaggctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181703 AGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGACAAGCTTGTTGAAGGCTT 181644

Qy    1081 gtcggatttctccggcgtgcgacgccgctttgagttccacggtgctatcgagggcggcaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181643 GTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGAGGGCGGCAA 181584

Qy    1141 atttaatggcgctgctatttatgatgattacgcacaccacccaacggaagtaactgcagt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181583 ATTTAATGGCGCTGCTATTTATGATGATTACGCACACCACCCAACGGAAGTAACTGCAGT 181524

Qy    1201 gctcagcgctgcgcgcacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgtt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181523 GCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAGGGCCGTGTCATCGTCGCGTT 181464

Qy    1261 ccaaccacatttatactcacgcaccatagaattccaaaaggagttcgcgggggcactgtc 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181463 CCAACCACATTTATACTCACGCACCATAGAATTCCAAAAGGAGTTCGCGGGGGCACTGTC 181404

Qy    1321 actggcagacgctgccgtggtgcttgagatttacggagcgcgcgaacaaccggtggatgg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181403 ACTGGCAGACGCTGCCGTGGTGCTTGAGATTTACGGAGCGCGCGAACAACCGGTGGATGG 181344

Qy    1381 cgtgtcctcggaaatcatcaccgatgcgatgaccattccagtggtgtacgaacctaattt 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181343 CGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCAGTGGTGTACGAACCTAATTT 181284

Qy    1441 ctctgcagtcccagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181283 CTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCTAATGACATCGTGCTCACCAT 181224

Qy    1501 gggtgcaggttccgtgaccatgcttgctccagaaatcctggatcagctgcaaaacaatta 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181223 GGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTGGATCAGCTGCAAAACAATTA 181164

Qy    1561 ggacgtaagtgaacaaggcag 1581
          |||||||||||||||||||||
Db 181163 GGACGTAAGTGAACAAGGCAG 181143
```

```
RESULT   2
AB015023
LOCUS       AB015023      2291 bp    DNA             BCT       06-FEB-1999
DEFINITION  Corynebacterium glutamicum genes for MurC and FtsQ, complete cds.
ACCESSION   AB015023
VERSION     AB015023.1  GI:3868934
KEYWORDS    FtsQ; MurC.
SOURCE      Corynebacterium glutamicum DNA.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 2291)
  AUTHORS   Wachi,M.                                          1026
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-1998) to the DDBJ/EMBL/GenBank databases. Masaaki
            Wachi, Tokyo Institute of Technology, Department of Bioengineering;
            4259 Nagatsuta Midori-ku, Yokohama, Kanagawa 226-8501, Japan
            (E-mail:mwachi@bio.titech.ac.jp, Tel:81-45-924-5770,
            Fax:81-45-924-5820)
REFERENCE   2  (sites)
  AUTHORS   Wachi,M., Wijayarathna,C.D., Teraoka,H. and Nagai,K.
  TITLE     A murC gene from coryneform bacteria
  JOURNAL   Appl. Microbiol. Biotechnol. (1998) In press
FEATURES             Location/Qualifiers
```

```
     source          1. .2291
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            94. .1554
                     /gene="murC"
     CDS             94. .1554
                     /gene="murC"
                     /codon_start=1
                     /transl_table=11
                     /product="MurC"
                     /protein_id="BAA34293.1"
                     /db_xref="GI:3868935"
                     /translation="MTTPHLDSAQDIDLSRVHLIGIGGAGMSGVARILLARGKTVTGS
                     DAKDSRTLLPLRAVGATIAVGHAAENLELSGELPTVVVTSPAAIPQDNPELVRAREEG
                     IPVIRRSDLLGELLEGSTQVLIAGTHGKTSTTSMSVVAMQAAGMDPSFAIGGQLNKAG
                     TNAHHGTGEVFIAEADESDASLLRYKPNVAVVTNVEPDHLDFFKTPEAYFQVFDDFAG
                     RITPNGKLVVCLNDPHAAELGERSVRKGIKTVGYGTADAVQAHPEVPAMATIVDSQVV
                     AEGTRATINIDGQEVSVILQIPGDHMVLHGAAALLAGYLVGGDVDKLVEGLSDFSGVR
                     RRFEPHGAIEGGKPNGAAIYDDYAHHPTEVTAVLSAARTRVKAAGKGRVIVAFQPHLY
                     SRTMEFQKEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVYEPNFSAV
                     PERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            1582. .2250
                     /gene="ftsQ"
     CDS             1582. .2250
                     /gene="ftsQ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsQ"
                     /protein_id="BAA34294.1"
                     /db_xref="GI:3868936"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKAQDAQM
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKRDGQRWNISSPSM
                     VTVR"
BASE COUNT      549 a    612 c    632 g    498 t
ORIGIN
```

Query Match 98.5%; Score 1558; DB 1; Length 2291;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1564; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy       8 acgcctccacggtgattgcagacatgattgctgcaactatcaatagccaacacaactaaa 67
          |||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1 ACGCGTCCACGGTGATTGCAGACATGATTGCTGCAACTATCAAAAGCCAACACAACTAAA 60

Qy      68 acgaccagctcaacgcaaaggaatagtttaaaggtgaccactccacacttggattctgca 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACGACCAGCTCAACGCAAAGGAATAGTTTAAAGGTGACCACTCCACACTTGGATTCTGCA 120

Qy     128 caagatattgatctgtcccgcgtccacctcatcggtattggcggagccggaatgtctggc 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAAGATATTGATCTGTCCCGCGTCCACCTCATCGGTATTGGCGGAGCCGGAATGTCTGGC 180

Qy     188 gttgcccgaatcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcc 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTTGCCCGAATCCTGCTTGCCCGCGGTAAGACAGTCACTGGTTCCGATGCCAAAGATTCC 240

Qy     248 cgcaccttgcttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaa 307
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     241 CGCACCTTGCTTCCACTCCGCGCCGTGGGAGCCACCATCGCAGTGGGACACGCCGCGGAA 300

Qy     308 aaccttgagctttccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccg 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AACCTTGAGCTTTCCGGCGAACTTCCCACCGTCGTGGTGACCTCTTTTGCCGCCATTCCG 360

Qy     368 caagacaacccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctcc 427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAAGACAACCCGGAACTTGTTCGTGCACGTGAAGAAGGCATTCCGGTTATTCGTCGCTCC 420

Qy     428 gatctgttgggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacggt 487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GATCTGTTGGGCGAATTGCTGGAAGGCTCCACCCAGGTCTTGATCGCGGGTACCCACGGT 480

Qy     488 aagacctccaccacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagc 547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAGACCTCCACCACCTCTATGTCTGTGGTAGCTATGCAGGCAGCGGGCATGGATCCAAGC 540

Qy     548 tttgctatcggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgag 607
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TTTGCTATCGGCGGACAGCTCAACAAGGCTGGCACCAATGCGCACCATGGAACTGGTGAG 600

Qy     608 gtctttatcgctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgtt 667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTCTTTATCGCTGAAGCAGATGAATCTGACGCATCGCTGCTGCGCTACAAGCCAAATGTT 660

Qy     668 gcagtggtcaccaatgtggaaccagaccacctggacttctttaaaaccccctgaagcctac 727
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCAGTGGTCACCAATGTGGAACCAGACCACCTGGACTTCTTTAAAACCCCTGAAGCCTAC 720

Qy     728 ttccaagtgttcgacgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgc 787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TTCCAAGTGTTCGACGATTTCGCAGGACGCATCACCCCGAACGGCAAGCTGGTTGTGTGC 780

Qy     788 ctgaacgatcctcacgcagcggagctgggggagaggtctgtccgcaagggtatcaagact 847
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTGAACGATCCTCACGCAGCGGAGCTGGGGGAGAGGTCTGTCCGCAAGGGTATCAAGACT 840

Qy     848 gttggttacggtaccgctgacgccgtacaggcacaccctgaggttccagcgatggctacc 907
          |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GTTGGTTATGGTACCGCTGACGCAGTACAGGCACACCCTGAGGTTCCAGCGATGGCTACC 900

Qy     908 atcgtggattcccaagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacag 967
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ATCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCATCAACATCGATGGACAG 960

Qy     968 gaagtatctgtgattcttcaaatccctggtgatcacatggtactcaacggtgcagccgcc 1027
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACTCAACGGTGCAGCCGCC 1020

Qy    1028 ctgctggccggatacctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggat 1087
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGACAAGCTTGTTGAAGGCTTGTCGGAT 1080

Qy    1088 ttctccggcgtgcgacgccgctttgagttccacggtgctatcgagggcggcaaatttaat 1147
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGAGGGCGGCAAATTTAAT 1140

Qy    1148 ggcgctgctatttatgatgattacgcacaccacccaacggaagtaactgcagtgctcagc 1207
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GGCGCGGCTATTTATGATGATTACGCACACCACCCAACGGAAGTAACTGCAGTGCTCAGC 1200

Qy    1208 gctgcgcgcacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaacca 1267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAGGGCCGTGTCATCGTCGCGTTCCAACCA 1260

Qy    1268 catttatactcacgcaccatagaattccaaaaggagttcgcgggggcactgtcactggca 1327
          ||||| |||||||||||||| ||||||||||||||||| ||||||||||||||||||||||
Db    1261 CATTTGTACTCACGCACCATGGAATTCCAAAAGGAGTTCGCGGAGGCACTGTCACTGGCA 1320

Qy    1328 gacgctgccgtggtgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcc 1387
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GACGCTGCCGTGGTGCTCGAGATTTACGGAGCGCGCGAACAACCGGTGGATGGCGTGTCC 1380

Qy    1388 tcggaaatcatcaccgatgcgatgaccattccagtggtgtacgaacctaatttctctgca 1447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TCGGAAATCATCACCGATGCGATGACCATTCCAGTGGTGTACGAACCTAATTTCTCTGCA 1440

Qy    1448 gtcccagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgca 1507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GTCCCAGAACGCATTGCAGAAATCGCAGGACCTAATGACATCGTGCTCACCATGGGTGCA 1500

Qy    1508 ggttccgtgaccatgcttgctccagaaatcctggatcagctgcaaaacaattaggacgta 1567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GGTTCCGTGACCATGCTTGCTCCAGAAATCCTGGATCAGCTGCAAAACAATTAGGACGTA 1560
```

```
Qy   1568 agtgaacaaggcag 1581
          ||||||||||||||
Db   1561 AGTGAACAAGGCAG 1574

RESULT   3
BLA242646
LOCUS       BLA242646    4622 bp    DNA             BCT       19-JAN-2000
DEFINITION  Brevibacterium lactofermentum murD (partial), ftsW, murG and murC
            genes, located upstream of ftsQ-ftsZ.
ACCESSION   AJ242646
VERSION     AJ242646.1  GI:6723455
KEYWORDS    ftsW gene; FtsW protein; murC gene; MurC protein; murD gene; MurD
            protein; murG gene; MurG protein.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 4622)
  AUTHORS   Ramos,A., Honrubia,P. and Gil,J.A.
  TITLE     Characterization and chromosomal organization of the murD-murC
            region of Brevibacterium lactofermentum ATCC 13869
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4622)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-1999) Gil J.A., Microbiology, University of Leon,
            Campus de Vegazana s/n, 24071 Leon, SPAIN
FEATURES             Location/Qualifiers
     source          1. .4622
                     /organism="Corynebacterium glutamicum"
                     /strain="ATCC 13869"
                     /db_xref="taxon:1718"
     gene            1. .333
                     /gene="murD"
     CDS             <1. .333
                     /gene="murD"
                     /function="peptidoglycan biosynthesis"
                     /codon_start=1
                     /transl_table=11
                     /evidence=experimental
                     /product="MurD protein"
                     /protein_id="CAB66322.1"
                     /db_xref="GI:6723456"
                     /translation="LKGADIAPLVKKHEQRIKAALVLGADRAEIVAALKEHAPQASVF
                     VTDKTEPFEAMEEIVTEAPSISEPGDTVLLAPAAASLDMFKGMGQRGDLFAHNIIGTI
                     KGLTEEKG"
     gene            336. .1808
                     /gene="ftsW"
     CDS             336. .1808
                     /gene="ftsW"
                     /function="cell division"
                     /codon_start=1
                     /transl_table=11
                     /evidence=experimental
                     /product="FtsW protein"
                     /protein_id="CAB66323.1"
                     /db_xref="GI:6723457"
                     /translation="MTTGASKKPARPNTGAKTRTGLGIRERISGAWNDLLARPLTDYI
                     MILCIVVILSCLGVVMVYSSSMTWSLREGGSVWGTAVRQGIMIVLGFFAMWVALMTRP
                     QTIRNLSNLILIVSIVLLLAVQIPGIGTGKEEVGSQSWIALGPIQFQPSEIAKVAIAV
                     WGAHYLAGKGPVQHWFNNHLMRFGGVGAFMAFLIFMEGDAGMAMSFVLVVLFMLFFAG
                     IAMGWIAIAGVLIIAALAVLAMGGGFRSSRFEVYFDALPGNFHDVRGIAFQSYQGFLS
                     LADGSGLGVGLGQSRAKWFYLPEAKNDFIFAIIGEELGLWGGALVIALPAGLLYFGLR
                     TAKKSHDPFLGLMAATLTASVVSQAFINIGYVVGLLPVTGIQLPMISAGGTSAIITLA
                     SMGLLISCARHEPETVSAMASYGRPAIDRLLGLREPSSTLTTSNASLRSNKTKAAKQR
                     SESSERVSGPLWRACDRTPSAGATKWASWSTIGSSATLDW"
     gene            1994. .3112
                     /gene="murG"
     CDS             1994. .3112
                     /gene="murG"
                     /function="peptidoglycan biosynthesis"
                     /codon_start=1
                     /transl_table=11
                     /product="MurG protein"
                     /protein_id="CAB66324.1"
                     /db_xref="GI:6723458"
                     /translation="MKSGELRIKMANSPKPMRVVVAGGGTAGHIEPALAVAEALRDKH
                     GATVSALGTARGLETTLVPDRGFELHLIEPVPVPRKPNMDLLKLPFRVAKALGQARKA
                     LKDTDAQAVIGFGGYVSRPAYMAAKSLGLPFFVHEANARAGMANKLGVKLGGVGLNAV
                     AGSGMDGDVVGIPIRAVLSGARDESAADRARDTWGLDKDRQTIFVTGGSQGSVSINKA
                     VEQAIDQLVEAGFQVLHAVGKKNELPAAKPGYHPVPPIDDMQAAYTVADLIVCRSGRM
                     TVAEVTAAGVPAIYVPLPHGNGEQALNAQAVIKAGAARQIDDADFTAQTLIDATLDIL
                     LHPSTRQSMSDAAKTSTAGNASTVIADMIAATINSQHN"
     gene            3147. .4607
                     /gene="murC"
     CDS             3147. .4607
                     /gene="murC"
                     /function="peptidoglycan biosynthesis"
                     /codon_start=1
                     /transl_table=11
                     /product="MurC protein"
                     /protein_id="CAB66325.1"
                     /db_xref="GI:6723459"
                     /translation="MTTPHLDSAQDIDLSRVHLIGIGGAGMSGVARILLARGKTVTGS
                     DAKDSRTLLPLRAVGATIAVGHAAENLELSGELPTVVVTSFAAIPQDNPELVRAREEG
                     IPVIRRSDLLGELLEGSTQVLIAGTHGKTSTTSMSVVAMQAAGMDPSFAIGGQLNKAG
                     TNAHHGTGEVFIAEADESDASLLRYKPNVAVVTNVEPDHLDFFKTPEAYFQVFDDPAG
                     RITPNGKLVVCLNDPHAAELGERSVRKGIKTVGYGTRDAVQAHPEVPAMATVVDSQVV
                     AEGTRATINIDGQEVSVILQIPGDHMVLNGAAALLAGYLVGGDVDKLVEGLSDPSGVR
                     RRFEPHGAIEGGKPNGAAIYDDYAHHPTEVTAVLRRCATRVKAAGKGRVIVAFQPHLY
                     SRTMEFQKEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVYEPNFSAV
                     PERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
BASE COUNT      989 a   1254 c   1317 g   1062 t
ORIGIN
```

Query Match 97.8%; Score 1545.6; DB 1; Length 4622;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1557; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy      1 gcaggtaacgcctccacggtgattgcagacatgattgctgcaactatcaatagccaacac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3047 GCAGGTAACGCCTCCACGGTGATTGCAGACATGATTGCTGCAACTATCAATAGCCAACAC 3106

Qy     61 aactaaaacgaccagctcaacgcaaaggaatagtttaaaggtgaccactccacacttgga 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3107 AACTAAAACGACCAGCTCAACGCAAAGGAATAGTTTAAAGGTGACCACTCCACACTTGGA 3166

Qy    121 ttctgcacaagatattgatctgtcccgcgtccacctcatcggtattggcggagccggaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3167 TTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATCGGTATTGGCGGAGCCGGAAT 3226

Qy    181 gtctggcgttgcccgaatcctgcttgcccgcggtaagacagtcactggttccgatgccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3227 GTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACAGTCACTGGTTCCGATGCCAA 3286

Qy    241 agattcccgcaccttgcttccactccgcgccgtgggagccaccatcgcagtgggacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3287 AGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCCACCATCGCAGTGGGACACGC 3346

Qy    301 tgcggaaaaccttgagctttccggcgaacttcccaccgtcgtggtgacctcttttgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3347 TGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTCGTGGTGACCTCTTTTGCCGC 3406

Qy    361 cattccgcaagacaacccggaacttgttcgtgcacgtgaagaaggcattccggttattcg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3407 CATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAAGAAGGCATTCCGGTTATTCG 3466

Qy    421 tcgctccgatctgttgggcgaattgctggaaggctccacccaggtcttgatcgcgggtac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3467 TCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACCCAGGTCTTGATCGCGGGTAC 3526

Qy    481 ccacggtaagacctccaccacctctatgtctgtggtagctatgcaggcagcgggcatgga 540
```

```
Db   3527 CCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCTATGCAGGCAGCGGGCATGGA 3586

Qy    541 tccaagctttgctatcggcggacagctcaacaaggctggcaccaatgcgcaccatggaac 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3587 TCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGCACCAATGCGCACCATGGAAC 3646

Qy    601 tggtgaggtctttatcgctgaagcagatgaatctgacgcatcgctgctgcgctacaagcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3647 TGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCATCGCTGCTGCGCTACAAGCC 3706

Qy    661 aaatgttgcagtggtcaccaatgtggaaccagaccacctggacttctttaaaaccccctga 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3707 AAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTGGACTTCTTTAAAACCCCTGA 3766

Qy    721 agcctacttccaagtgttcgacgatttcgcaggacgcatcaccccgaacggcaagctggt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3767 AGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATCACCCCGAACGGCAAGCTGGT 3826

Qy    781 tgtgtgcctgaacgatcctcacgcagcggagctgggggagaggtctgtccgcaagggtat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3827 TGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAGAGGTCTGTCCGCAAGGGTAT 3886

Qy    841 caagactgttggttacggtaccgctgacgccgtacaggcacaccctgaggttccagcgat 900
          ||||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||
Db   3887 CAAGACTGTTGGTTACGGTACCCGTGACGCAGTACAGGCACACCCTGAGGTTCCAGCGAT 3946

Qy    901 ggctaccatcgtggattcccaagttgtcgcagaaggcacccgcgccaccatcaacatcga 960
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3947 GGCTACCGTCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCATCAACATCGA 4006

Qy    961 tggacaggaagtatctgtgattcttcaaatccctggtgatcacatggtactcaacggtgc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4007 TGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACTCAACGGTGC 4066

Qy   1021 agccgccctgctggccggatacctggtgggtggggacgtcgacaagcttgttgaaggctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4067 AGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGATAAGCTTGTTGAAGGCTT 4126

Qy   1081 gtcggatttctccggcgtgcgacgccgctttgagttccacggtgctatcgagggcggcaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4127 GTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGAGGGCGGCAA 4186

Qy   1141 atttaatggcgctgctatttatgatgattacgcacaccacccaacggaagtaactgcagt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4187 ATTTAATGGCGCTGCTATTTATGATGATTACGCACACCACCCAACGGAAGTAACTGCAGT 4246

Qy   1201 gctcagcgctgcgcgcacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgtt 1260
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4247 GCTCAGGCGCTGCGCGACCCGGGTGAAGGCCGCTGGAAAGGGCCGTGTCATCGTCGCGTT 4306

Qy   1261 ccaaccacatttatactcacgcaccatagaattccaaaaggagttcgcgggggcactgtc 1320
          |||||||||||| |||||||||||||| |||||||||||||||||||| |||||||||
Db   4307 CCAACCACATTTGTACTCACGCACCATGGAATTCCAAAAGGAGTTCGCGGAGGCACTGTC 4366

Qy   1321 actggcagacgctgccgtggtgcttgagatttacggagcgcgcgaacaaccggtggatgg 1380
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   4367 ACTGGCAGACGCTGCCGTGGTGCTCGAGATTTACGGAGCGCGCGAACAACCGGTGGATGG 4426

Qy   1381 cgtgtcctcggaaatcatcaccgatgcgatgaccattccagtggtgtacgaacctaattt 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4427 CGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCAGTGGTGTACGAACCTAATTT 4486

Qy   1441 ctctgcagtcccagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4487 CTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCTAATGACATCGTGCTCACCAT 4546

Qy   1501 gggtgcaggttccgtgaccatgcttgctccagaaatcctggatcagctgcaaaacaatta 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4547 GGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTGGATCAGCTGCAAAACAATTA 4606

Qy   1561 ggacgtaagtgaacaa 1576
          ||||||||||||||||
```

```
Db   4607 GGACGTAAGTGAACAA 4622

RESULT   4
AX122446
LOCUS       AX122446     1458 bp    DNA             PAT       11-MAY-2001
DEFINITION  Sequence 2362 from Patent EP1108790.
ACCESSION   AX122446
VERSION     AX122446.1  GI:14039691
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 1458)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 2362 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1..1458
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
BASE COUNT      318 a    420 c    400 g    320 t
ORIGIN

  Query Match           92.2%;  Score 1458;  DB 6;  Length 1458;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1458;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    101 gtgaccactccacacttggattctgcacaagatattgatctgtcccgcgtccacctcatc 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACCACTCCACACTTGGATTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATC 60

Qy    161 ggtattggcggagccggaatgtctggcgttgcccgaatcctgcttgcccgcggtaagaca 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTATTGGCGGAGCCGGAATGTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACA 120

Qy    221 gtcactggttccgatgccaaagattcccgcaccttgcttccactccgcgccgtgggagcc 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTCACTGGTTCCGATGCCAAAGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCC 180

Qy    281 accatcgcagtgggacacgctgcggaaaaccttgagctttccggcgaacttcccaccgtc 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCATCGCAGTGGGACACGCTGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTC 240

Qy    341 gtggtgacctcttttgccgccattccgcaagacaacccggaacttgttcgtgcacgtgaa 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGGTGACCTCTTTTGCCGCCATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAA 300

Qy    401 gaaggcattccggttattcgtcgctccgatctgttgggcgaattgctggaaggctccacc 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAAGGCATTCCGGTTATTCGTCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACC 360

Qy    461 caggtcttgatcgcgggtacccacggtaagacctccaccacctctatgtctgtggtagct 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAGGTCTTGATCGCGGGTACCCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCT 420

Qy    521 atgcaggcagcgggcatggatccaagctttgctatcggcggacagctcaacaaggctggc 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGCAGGCAGCGGGCATGGATCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGC 480

Qy    581 accaatgcgcaccatggaactggtgaggtctttatcgctgaagcagatgaatctgacgca 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCAATGCGCACCATGGAACTGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCA 540

Qy    641 tcgctgctgcgctacaagccaaatgttgcagtggtcaccaatgtggaaccagaccacctg 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCGCTGCTGCGCTACAAGCCAAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTG 600
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 09:58:11 ; Search time 1598.42 Seconds
(without alignments)
16317.389 Million cell updates/sec

Title: US-09-603-124B-1
Perfect score: 1581
Sequence: 1 gcaggtaacgcctccacggt.....gacgtaagtgaacaaggcag 1581

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
c	1	1581	100.0	349980	6	AX127150	AX127150 Sequence
	2	1458	92.2	1458	6	AX122446	AX122446 Sequence
	3	1122	71.0	2291	1	AB015023	AB015023 Corynebac
	4	1050	66.4	1458	6	E28467	E28467 Novel murC
	5	951	60.2	4622	1	BLA242646	AJ242646 Brevibact
	6	236	14.9	4116	1	AB003132	AB003132 Corynebac
	7	236	14.9	5546	1	BLFTSZ	Y08964 Brevibacter
	8	63	4.0	1116	6	AX122447	AX122447 Sequence
	9	21	1.3	1073	8	AF022730	AF022730 Oryza sat
c	10	21	1.3	6843	2	AC019805	AC019805 Drosophil
	11	21	1.3	85875	2	AC010010	AC010010 Drosophil
c	12	21	1.3	126039	9	AC012665	AC012665 Homo sapi
c	13	21	1.3	137569	2	AC084093	AC084093 Homo sapi
	14	21	1.3	163027	2	AC022846	AC022846 Homo sapi
	15	21	1.3	174822	9	AC013275	AC013275 Homo sapi
c	16	21	1.3	175590	2	AF307159	AF307159 Homo sapi
c	17	21	1.3	181848	2	AC090199	AC090199 Homo sapi
c	18	21	1.3	310120	3	AE003468	AE003468 Drosophil
	19	20	1.3	234	8	TAE131670	AJ131670 Triticum
c	20	20	1.3	1263	6	AX122173	AX122173 Sequence
	21	20	1.3	1374	6	AX122688	AX122688 Sequence
	22	20	1.3	1556	6	AX064415	AX064415 Sequence
	23	20	1.3	52650	2	AC010026	AC010026 Drosophil
c	24	20	1.3	157541	2	AC008523	AC008523 Homo sapi
c	25	20	1.3	349980	6	AX127149	AX127149 Sequence
c	26	20	1.3	349980	6	AX127151	AX127151 Sequence
	27	19	1.2	11994	1	AE005031	AE005031 Halobacte
c	28	19	1.2	15000	8	AF169016	AF169016 Aspergill
	29	19	1.2	32274	1	SCC8A	AL356892 Streptomy
	30	19	1.2	56917	1	AME16952	Y16952 Amycolatops
	31	19	1.2	67480	2	AF279874	AF279874 Homo sapi
c	32	19	1.2	85217	8	ATT17J13	AL138651 Arabidops
	33	19	1.2	104770	9	HSJ245M18	AL078584 Human DNA
	34	19	1.2	109612	9	AL512503	AL512503 Human DNA
	35	19	1.2	114856	9	HSDJ80E14	AL117375 Human DNA
c	36	19	1.2	116501	2	AC008053	AC008053 Homo sapi
	37	19	1.2	119060	2	AC023390	AC023390 Homo sapi
c	38	19	1.2	119456	2	AC084789	AC084789 Homo sapi
	39	19	1.2	131047	2	AC002421	AC002421 Homo sapi
	40	19	1.2	139232	2	AL353762	AL353762 Homo sapi
	41	19	1.2	144727	9	HSJ646P11	AL049715 Human DNA
	42	19	1.2	145382	2	AC073223	AC073223 Homo sapi
	43	19	1.2	148260	2	AL596222	AL596222 Homo sapi
	44	19	1.2	148882	2	AC018381	AC018381 Homo sapi
c	45	19	1.2	150436	2	AC024119	AC024119 Homo sapi

ALIGNMENTS

RESULT 1
AX127150/c
LOCUS AX127150 349980 bp DNA PAT 11-MAY-2001
DEFINITION Sequence 7066 from Patent EP1108790.
ACCESSION AX127150 AX114121
VERSION AX127150.1 GI:14041138
KEYWORDS .
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Corynebacteriaceae;
Corynebacterium.
REFERENCE 1 (bases 1 to 349980)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7066 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES Location/Qualifiers
source 1. .349980

```
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
                     /note="Seq 1 to long (3.309.400) split in 11, seq 7066
                     2.100.001 2.449.980"
BASE COUNT     80724 a   98367 c   90490 g   80399 t
ORIGIN

  Query Match           100.0%;  Score 1581;  DB 6;  Length 349980;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1581;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gcaggtaacgcctccacggtgattgcagacatgattgctgcaactatcaatagccaacac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182723 GCAGGTAACGCCTCCACGGTGATTGCAGACATGATTGCTGCAACTATCAATAGCCAACAC 182664

Qy      61 aactaaaacgaccagctcaacgcaaaggaatagtttaaaggtgaccactccacacttgga 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182663 AACTAAAACGACCAGCTCAACGCAAAGGAATAGTTTAAAGGTGACCACTCCACACTTGGA 182604

Qy     121 ttctgcacaagatattgatctgtcccgcgtccacctcatcggtattggcggagccggaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182603 TTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATCGGTATTGGCGGAGCCGGAAT 182544

Qy     181 gtctggcgttgcccgaatcctgcttgcccgcggtaagacagtcactggttccgatgccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182543 GTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACAGTCACTGGTTCCGATGCCAA 182484

Qy     241 agattcccgcaccttgcttccactccgcgccgtgggagccaccatcgcagtgggacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182483 AGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCCACCATCGCAGTGGGACACGC 182424

Qy     301 tgcggaaaaccttgagctttccggcgaacttcccaccgtcgtggtgacctcttttgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182423 TGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTCGTGGTGACCTCTTTTGCCGC 182364

Qy     361 cattccgcaagacaacccggaacttgttcgtgcacgtgaagaaggcattccggttattcg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182363 CATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAAGAAGGCATTCCGGTTATTCG 182304

Qy     421 tcgctccgatctgttgggcgaattgctggaaggctccacccaggtcttgatcgcgggtac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182303 TCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACCCAGGTCTTGATCGCGGGTAC 182244

Qy     481 ccacggtaagacctccaccacctctatgtctgtggtagctatgcaggcagcgggcatgga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182243 CCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCTATGCAGGCAGCGGGCATGGA 182184

Qy     541 tccaagctttgctatcggcggacagctcaacaaggctggcaccaatgcgcaccatggaac 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182183 TCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGCACCAATGCGCACCATGGAAC 182124

Qy     601 tggtgaggtctttatcgctgaagcagatgaatctgacgcatcgctgctgcgctacaagcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182123 TGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCATCGCTGCTGCGCTACAAGCC 182064

Qy     661 aaatgttgcagtggtcaccaatgtggaaccagaccacctggacttctttaaaaccccctga 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182063 AAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTGGACTTCTTTAAAACCCCTGA 182004

Qy     721 agcctacttccaagtgttcgacgatttcgcaggacgcatcaccccgaacggcaagctggt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 182003 AGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATCACCCCGAACGGCAAGCTGGT 181944

Qy     781 tgtgtgcctgaacgatcctcacgcagcggagctgggggagaggtctgtccgcaagggtat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181943 TGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAGAGGTCTGTCCGCAAGGGTAT 181884

Qy     841 caagactgttggttacggtaccgctgacgccgtacaggcacaccctgaggttccagcgat 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181883 CAAGACTGTTGGTTACGGTACCGCTGACGCCGTACAGGCACACCCTGAGGTTCCAGCGAT 181824

Qy     901 ggctaccatcgtggattcccaagttgtcgcagaaggcacccgcgccaccatcaacatcga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181823 GGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCATCAACATCGA 181764

Qy     961 tggacaggaagtatctgtgattcttcaaatccctggtgatcacatggtactcaacggtgc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181763 TGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACTCAACGGTGC 181704

Qy    1021 agccgccctgctggccggatacctggtgggtggggacgtcgacaagcttgttgaaggctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181703 AGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGACAAGCTTGTTGAAGGCTT 181644

Qy    1081 gtcggatttctccggcgtgcgacgccgctttgagttccacggtgctatcgagggcggcaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181643 GTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGAGGGCGGCAA 181584

Qy    1141 atttaatggcgctgctatttatgatgattacgcacaccacccaacggaagtaactgcagt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181583 ATTTAATGGCGCTGCTATTTATGATGATTACGCACACCACCCAACGGAAGTAACTGCAGT 181524

Qy    1201 gctcagcgctgcgcgcacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgtt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181523 GCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAGGGCCGTGTCATCGTCGCGTT 181464

Qy    1261 ccaaccacatttatactcacgcaccatagaattccaaaaggagttcgcgggggcactgtc 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181463 CCAACCACATTTATACTCACGCACCATAGAATTCCAAAAGGAGTTCGCGGGGGCACTGTC 181404

Qy    1321 actggcagacgctgccgtggtgcttgagatttacggagcgcgcgaacaaccggtggatgg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181403 ACTGGCAGACGCTGCCGTGGTGCTTGAGATTTACGGAGCGCGCGAACAACCGGTGGATGG 181344

Qy    1381 cgtgtcctcggaaatcatcaccgatgcgatgaccattccagtggtgtacgaacctaattt 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181343 CGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCAGTGGTGTACGAACCTAATTT 181284

Qy    1441 ctctgcagtcccagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccat 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181283 CTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCTAATGACATCGTGCTCACCAT 181224

Qy    1501 gggtgcaggttccgtgaccatgcttgctccagaaatcctggatcagctgcaaaacaatta 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181223 GGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTGGATCAGCTGCAAAACAATTA 181164

Qy    1561 ggacgtaagtgaacaaggcag 1581
          |||||||||||||||||||||
Db 181163 GGACGTAAGTGAACAAGGCAG 181143


RESULT  2
AX122446
LOCUS       AX122446     1458 bp    DNA             PAT       11-MAY-2001
DEFINITION  Sequence 2362 from Patent EP1108790.
ACCESSION   AX122446
VERSION     AX122446.1  GI:14039691
KEYWORDS
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 1458)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 2362 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1. .1458
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
BASE COUNT      318 a    420 c    400 g    320 t
ORIGIN
```

Query Match 92.2%; Score 1458; DB 6; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    101 gtgaccactccacacttggattctgcacaagatattgatctgtcccgcgtccacctcatc 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACCACTCCACACTTGGATTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATC 60

Qy    161 ggtattggcggagccggaatgtctggcgttgcccgaatcctgcttgcccgcggtaagaca 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTATTGGCGGAGCCGGAATGTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACA 120

Qy    221 gtcactggttccgatgccaaagattcccgcaccttgcttccactccgcgccgtgggagcc 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTCACTGGTTCCGATGCCAAAGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCC 180

Qy    281 accatcgcagtgggacacgctgcggaaaaccttgagctttccggcgaacttcccaccgtc 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCATCGCAGTGGGACACGCTGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTC 240

Qy    341 gtggtgacctcttttgccgccattccgcaagacaacccggaacttgttcgtgcacgtgaa 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGGTGACCTCTTTTGCCGCCATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAA 300

Qy    401 gaaggcattccggttattcgtcgctccgatctgttgggcgaattgctggaaggctccacc 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAAGGCATTCCGGTTATTCGTCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACC 360

Qy    461 caggtcttgatcgcgggtacccacggtaagacctccaccacctctatgtctgtggtagct 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAGGTCTTGATCGCGGGTACCCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCT 420

Qy    521 atgcaggcagcgggcatggatccaagctttgctatcggcggacagctcaacaaggctggc 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGCAGGCAGCGGGCATGGATCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGC 480

Qy    581 accaatgcgcaccatggaactggtgaggtctttatcgctgaagcagatgaatctgacgca 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCAATGCGCACCATGGAACTGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCA 540

Qy    641 tcgctgctgcgctacaagccaaatgttgcagtggtcaccaatgtggaaccagaccacctg 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCGCTGCTGCGCTACAAGCCAAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTG 600

Qy    701 gacttctttaaaaccccctgaagcctacttccaagtgttcgacgatttcgcaggacgcatc 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACTTCTTTAAAACCCCTGAAGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATC 660

Qy    761 accccgaacggcaagctggttgtgtgcctgaacgatcctcacgcagcggagctgggggag 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCCGAACGGCAAGCTGGTTGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAG 720

Qy    821 aggtctgtccgcaagggtatcaagactgttggttacggtaccgctgacgccgtacaggca 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGGTCTGTCCGCAAGGGTATCAAGACTGTTGGTTACGGTACCGCTGACGCCGTACAGGCA 780

Qy    881 caccctgaggttccagcgatggctaccatcgtggattcccaagttgtcgcagaaggcacc 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACCCTGAGGTTCCAGCGATGGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACC 840

Qy    941 cgcgccaccatcaacatcgatggacaggaagtatctgtgattcttcaaatccctggtgat 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCGCCACCATCAACATCGATGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGAT 900

Qy   1001 cacatggtactcaacggtgcagccgccctgctggccggatacctggtgggtggggacgtc 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CACATGGTACTCAACGGTGCAGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTC 960

Qy   1061 gacaagcttgttgaaggcttgtcggatttctccggcgtgcgacgccgctttgagttccac 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACAAGCTTGTTGAAGGCTTGTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCAC 1020

Qy   1121 ggtgctatcgagggcggcaaatttaatggcgctgctatttatgatgattacgcacaccac 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGTGCTATCGAGGGCGGCAAATTTAATGGCGCTGCTATTTATGATGATTACGCACACCAC 1080

Qy   1181 ccaacggaagtaactgcagtgctcagcgctgcgcgcacccgggtgaaggccgctggaaag 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCAACGGAAGTAACTGCAGTGCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAG 1140

Qy   1241 ggccgtgtcatcgtcgcgttccaaccacatttatactcacgcaccatagaattccaaaag 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCCGTGTCATCGTCGCGTTCCAACCACATTTATACTCACGCACCATAGAATTCCAAAAG 1200

Qy   1301 gagttcgcgggggcactgtcactggcagacgctgccgtggtgcttgagatttacggagcg 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGTTCGCGGGGGCACTGTCACTGGCAGACGCTGCCGTGGTGCTTGAGATTTACGGAGCG 1260

Qy   1361 cgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgatgcgatgaccattcca 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCA 1320

Qy   1421 gtggtgtacgaacctaatttctctgcagtcccagaacgcattgcagaaatcgcaggacct 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCT 1380

Qy   1481 aatgacatcgtgctcaccatgggtgcaggttccgtgaccatgcttgctccagaaatcctg 1540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTG 1440

Qy   1541 gatcagctgcaaaacaat 1558
          ||||||||||||||||||
Db   1441 GATCAGCTGCAAAACAAT 1458
```

```
RESULT   3
AB015023
LOCUS       AB015023     2291 bp    DNA             BCT       06-FEB-1999
DEFINITION  Corynebacterium glutamicum genes for MurC and FtsQ, complete cds.
ACCESSION   AB015023
VERSION     AB015023.1  GI:3868934
KEYWORDS    FtsQ; MurC.
SOURCE      Corynebacterium glutamicum DNA.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 2291)
  AUTHORS   Wachi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-1998) to the DDBJ/EMBL/GenBank databases. Masaaki
            Wachi, Tokyo Institute of Technology, Department of Bioengineering;
            4259 Nagatsuta Midori-ku, Yokohama, Kanagawa 226-8501, Japan
            (E-mail:mwachi@bio.titech.ac.jp, Tel:81-45-924-5770,
            Fax:81-45-924-5820)
REFERENCE   2  (sites)
  AUTHORS   Wachi,M., Wijayarathna,C.D., Teraoka,H. and Nagai,K.
  TITLE     A murC gene from coryneform bacteria
  JOURNAL   Appl. Microbiol. Biotechnol. (1998) In press
FEATURES             Location/Qualifiers
     source          1. .2291
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            94. .1554
                     /gene="murC"
     CDS             94. .1554
                     /gene="murC"
                     /codon_start=1
                     /transl_table=11
                     /product="MurC"
                     /protein_id="BAA34293.1"
                     /db_xref="GI:3868935"
                     /translation="MTTPHLDSAQDIDLSRVHLIGIGGAGMSGVARILLARGKTVTGS
                     DAKDSRTLLPLRAVGATIAVGHAAENLELSGELPTVVVTSFAAIPQDNPELVRAREEG
                     IPVIRRSDLLGELLEGSTQVLIAGTHGKTSTTSMSVVAMQAAGMDPSFAIGGQLNKAG
```

```
                     TNAHHGTGEVFIAEADESDASLLRYKPNVAVVTNVEPDHLDFFKTPEAYFQVFDDFAG
                     RITPNGKLVVCLNDPHAAELGERSVRKGIKTVGYGTADAVQAHPEVPAMATIVDSQVV
                     AEGTRATINIDGQEVSVILQIPGDHMVLNGAAALLAGYLVGGDVDKLVEGLSDFSGVR
                     RRFEFHGAIEGGKFNGAAIYDDYAHHPTEVTAVLSAARTRVKAAGKGRVIVAFQPHLY
                     SRTMEFQKEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVYEPNFSAV
                     PERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            1582. .2250
                     /gene="ftsQ"
     CDS             1582. .2250
                     /gene="ftsQ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsQ"
                     /protein_id="BAA34294.1"
                     /db_xref="GI:3868936"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLPRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKAQDAQM
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKRDGQRWNISSPSM
                     VTVR"
BASE COUNT      549 a    612 c    632 g    498 t
ORIGIN

  Query Match             71.0%;  Score 1122;  DB 1;  Length 2291;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1522;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         52 agccaacacaactaaaacgaccagctcaacgcaaaggaatagtttaaaggtgaccactcc 111
Db         45 AGCCAACACAACTAAAACGACCAGCTCAACGCAAAGGAATAGTTTAAAGGTGACCACTCC 104

Qy        112 acacttggattctgcacaagatattgatctgtcccgcgtccacctcatcggtattggcgg 171
Db        105 ACACTTGGATTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATCGGTATTGGCGG 164

Qy        172 agccggaatgtctggcgttgcccgaatcctgcttgcccgcggtaagacagtcactggttc 231
Db        165 AGCCGGAATGTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACAGTCACTGGTTC 224

Qy        232 cgatgccaaagattcccgcaccttgcttccactccgcgccgtgggagccaccatcgcagt 291
Db        225 CGATGCCAAAGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCCACCATCGCAGT 284

Qy        292 gggacacgctgcggaaaaccttgagctttccggcgaacttcccaccgtcgtggtgacctc 351
Db        285 GGGACACGCCGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTCGTGGTGACCTC 344

Qy        352 ttttgccgccattccgcaagacaacccggaacttgttcgtgcacgtgaagaaggcattcc 411
Db        345 TTTTGCCGCCATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAAGAAGGCATTCC 404

Qy        412 ggttattcgtcgctccgatctgttgggcgaattgctggaaggctccacccaggtcttgat 471
Db        405 GGTTATTCGTCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACCCAGGTCTTGAT 464

Qy        472 cgcgggtacccacggtaagacctccaccacctctatgtctgtggtagctatgcaggcagc 531
Db        465 CGCGGGTACCCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCTATGCAGGCAGC 524

Qy        532 gggcatggatccaagctttgctatcggcggacagctcaacaaggctggcaccaatgcgca 591
Db        525 GGGCATGGATCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGCACCAATGCGCA 584

Qy        592 ccatggaactggtgaggtctttatcgctgaagcagatgaatctgacgcatcgctgctgcg 651
Db        585 CCATGGAACTGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCATCGCTGCTGCG 644

Qy        652 ctacaagccaaatgttgcagtggtcaccaatgtggaaccagaccacctggacttctttaa 711
Db        645 CTACAAGCCAAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTGGACTTCTTTAA 704

Qy        712 aacccctgaagcctacttccaagtgttcgacgatttcgcaggacgcatcaccccgaacgg 771
Db        705 AACCCCTGAAGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATCACCCCGAACGG 764

Qy        772 caagctggttgtgtgcctgaacgatcctcacgcagcggagctgggggagaggtctgtccg 831
Db        765 CAAGCTGGTTGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAGAGGTCTGTCCG 824

Qy        832 caagggtatcaagactgttggttacggtaccgctgacgccgtacaggcacaccctgaggt 891
Db        825 CAAGGGTATCAAGACTGTTGGTTATGGTACCGCTGACGCAGTACAGGCACACCCTGAGGT 884

Qy        892 tccagcgatggctaccatcgtggattcccaagttgtcgcagaaggcacccgcgccaccat 951
Db        885 TCCAGCGATGGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCAT 944

Qy        952 caacatcgatggacaggaagtatctgtgattcttcaaatccctggtgatcacatggtact 1011
Db        945 CAACATCGATGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACT 1004

Qy       1012 caacggtgcagccgccctgctggccggatacctggtgggtggggacgtcgacaagcttgt 1071
Db       1005 CAACGGTGCAGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGACAAGCTTGT 1064

Qy       1072 tgaaggcttgtcggatttctccggcgtgcgacgccgctttgagttccacggtgctatcga 1131
Db       1065 TGAAGGCTTGTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGA 1124

Qy       1132 gggcggcaaatttaatggcgctgctatttatgatgattacgcacaccacccaacggaagt 1191
Db       1125 GGGCGGCAAATTTAATGGCGCGGCTATTTATGATGATTACGCACACCACCCAACGGAAGT 1184

Qy       1192 aactgcagtgctcagcgctgcgcgcacccgggtgaaggccgctggaaagggccgtgtcat 1251
Db       1185 AACTGCAGTGCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAGGGCCGTGTCAT 1244

Qy       1252 cgtcgcgttccaaccacatttatactcacgcaccatagaattccaaaaggagttcgcggg 1311
Db       1245 CGTCGCGTTCCAACCACATTTGTACTCACGCACCATGGAATTCCAAAAGGAGTTCGCGGA 1304

Qy       1312 ggcactgtcactggcagacgctgccgtggtgcttgagatttacggagcgcgcgaacaacc 1371
Db       1305 GGCACTGTCACTGGCAGACGCTGCCGTGGTGCTCGAGATTTACGGAGCGCGCGAACAACC 1364

Qy       1372 ggtggatggcgtgtcctcggaaatcatcaccgatgcgatgaccattccagtggtgtacga 1431
Db       1365 GGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCAGTGGTGTACGA 1424

Qy       1432 acctaatttctctgcagtcccagaacgcattgcagaaatcgcaggacctaatgacatcgt 1491
Db       1425 ACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCTAATGACATCGT 1484

Qy       1492 gctcaccatgggtgcaggttccgtgaccatgcttgctccagaaatcctggatcagctgca 1551
Db       1485 GCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTGGATCAGCTGCA 1544

Qy       1552 aaacaattaggacgtaagtgaacaaggcag 1581
Db       1545 AAACAATTAGGACGTAAGTGAACAAGGCAG 1574

RESULT 4
E28467
LOCUS       E28467         1458 bp    DNA             PAT       07-FEB-2001
DEFINITION  Novel murC gene.
ACCESSION   E28467
VERSION     E28467.1  GI:13018359
KEYWORDS    JP 1999196876-A/1.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 1458)
```

```
  AUTHORS   Takashi,K.H.Y. and Wachi,K.N.
  TITLE     Novel murC gene
  JOURNAL   Patent: JP 1999196876-A 1 27-JUL-1999;
            MITSUBISHI CHEM CORP
COMMENT     OS   Brevibacterium flavum
            PN   JP 1999196876-A/1
            PD   27-JUL-1999;
            PP   16-JAN-1998 JP 1998006840
            PR
            PI   TAKASHI KOBAYASHI,HIDEAKI YUKAWA,MASAAKI WACHI,KAZUO NAGAI PC
             C12N15/09,C12N9/00//(C12N15/09,C12R1:13),C12N15/00,(C12N15/00, PC
             C12R1:13)
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   CDS             1. .1458.
FEATURES    Location/Qualifiers
     source     1. .1458
                /organism="Corynebacterium glutamicum"
                /db_xref="taxon:1718"
BASE COUNT      318 a    420 c    402 g    318 t
ORIGIN
```

102 A

Query Match 66.4%; Score 1050; DB 6; Length 1458;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1450; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy    101 gtgaccactccacacttggattctgcacaagatattgatctgtcccgcgtccacctcatc 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACCACTCCACACTTGGATTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATC 60

Qy    161 ggtattggcggagccggaatgtctggcgttgcccgaatcctgcttgcccgcggtaagaca 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTATTGGCGGAGCCGGAATGTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACA 120

Qy    221 gtcactggttccgatgccaaagattcccgcaccttgcttccactccgcgccgtgggagcc 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTCACTGGTTCCGATGCCAAAGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCC 180

Qy    281 accatcgcagtgggacacgctgcggaaaaccttgagctttccggcgaacttcccaccgtc 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCATCGCAGTGGGACACGCCGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTC 240

Qy    341 gtggtgacctcttttgccgccattccgcaagacaacccggaacttgttcgtgcacgtgaa 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGGTGACCTCTTTTGCCGCCATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAA 300

Qy    401 gaaggcattccggttattcgtcgctccgatctgttgggcgaattgctggaaggctccacc 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAAGGCATTCCGGTTATTCGTCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACC 360

Qy    461 caggtcttgatcgcgggtacccacggtaagacctccaccacctctatgtctgtggtagct 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAGGTCTTGATCGCGGGTACCCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCT 420

Qy    521 atgcaggcagcgggcatggatccaagctttgctatcggcggacagctcaacaaggctggc 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGCAGGCAGCGGGCATGGATCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGC 480

Qy    581 accaatgcgcaccatggaactggtgaggtctttatcgctgaagcagatgaatctgacgca 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCAATGCGCACCATGGAACTGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCA 540

Qy    641 tcgctgctgcgctacaagccaaatgttgcagtggtcaccaatgtggaaccagaccacctg 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCGCTGCTGCGCTACAAGCCAAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTG 600

Qy    701 gacttctttaaaaccccctgaagcctacttccaagtgttcgacgatttcgcaggacgcatc 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACTTCTTTAAAACCCCTGAAGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATC 660

Qy    761 accccgaacggcaagctggttgtgtgcctgaacgatcctcacgcagcggagctgggggag 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCCGAACGGCAAGCTGGTTGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAG 720

Qy    821 aggtctgtccgcaagggtatcaagactgttggttacggtaccgctgacgccgtacaggca 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGGTCTGTCCGCAAGGGTATCAAGACTGTTGGTTATGGTACCGCTGACGCAGTACAGGCA 780

Qy    881 caccctgaggttccagcgatggctaccatcgtggattcccaagttgtcgcagaaggcacc 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACCCTGAGGTTCCAGCGATGGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACC 840

Qy    941 cgcgccaccatcaacatcgatggacaggaagtatctgtgattcttcaaatccctggtgat 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCGCCACCATCAACATCGATGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGAT 900

Qy   1001 cacatggtactcaacggtgcagccgccctgctggccggatacctggtgggtggggacgtc 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CACATGGTACTCAACGGTGCAGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTC 960

Qy   1061 gacaagcttgttgaaggcttgtcggatttctccggcgtgcgacgccgctttgagttccac 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACAAGCTTGTTGAAGGCTTGTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCAC 1020

Qy   1121 ggtgctatcgagggcggcaaatttaatggcgctgctatttatgatgattacgcacaccac 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGTGCTATCGAGGGCGGCAAATTTAATGGCGCGGCTATTTATGATGATTACGCACACCAC 1080

Qy   1181 ccaacggaagtaactgcagtgctcagcgctgcgcgcacccgggtgaaggccgctggaaag 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCAACGGAAGTAACTGCAGTGCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAG 1140

Qy   1241 ggccgtgtcatcgtcgcgttccaaccacatttatactcacgcaccatagaattccaaaag 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCCGTGTCATCGTCGCGTTCCAACCACATTTGTACTCACGCACCATGGAATTCCAAAAG 1200

Qy   1301 gagttcgcgggggcactgtcactggcagacgctgccgtggtgcttgagatttacggagcg 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGTTCGCGGAGGCACTGTCACTGGCAGACGCTGCCGTGGTGCTCGAGATTTACGGAGCG 1260

Qy   1361 cgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgatgcgatgaccattcca 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCA 1320

Qy   1421 gtggtgtacgaacctaatttctctgcagtcccagaacgcattgcagaaatcgcaggacct 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCT 1380

Qy   1481 aatgacatcgtgctcaccatgggtgcaggttccgtgaccatgcttgctccagaaatcctg 1540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTG 1440

Qy   1541 gatcagctgcaaaacaat 1558
          ||||||||||||||||||
Db   1441 GATCAGCTGCAAAACAAT 1458
```

```
RESULT   5
BLA242646
LOCUS       BLA242646    4622 bp    DNA             BCT       19-JAN-2000
DEFINITION  Brevibacterium lactofermentum murD (partial), ftsW, murG and murC
            genes, located upstream of ftsQ-ftsZ.
ACCESSION   AJ242646
VERSION     AJ242646.1  GI:6723455
KEYWORDS    ftsW gene; FtsW protein; murC gene; MurC protein; murD gene; MurD
            protein; murG gene; MurG protein.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 4622)
  AUTHORS   Ramos,A., Honrubia,P. and Gil,J.A.
  TITLE     Characterization and chromosomal organization of the murD-murC
```

region of Brevibacterium lactofermentum ATCC 13869
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4622) 102A
AUTHORS Gil,J.A.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1999) Gil J.A., Microbiology, University of Leon, Campus de Vegazana s/n, 24071 Leon, SPAIN

FEATURES Location/Qualifiers

source 1. .4622
/organism="Corynebacterium glutamicum"
/strain="ATCC 13869"
/db_xref="taxon:1718"

gene 1. .333
/gene="murD"

CDS <1. .333
/gene="murD"
/function="peptidoglycan biosynthesis"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="MurD protein"
/protein_id="CAB66322.1"
/db_xref="GI:6723456"
/translation="LKGADIAPLVKKHEQRIKAALVLGADRAEIVAALKEHAPQASVF
VTDKTEPFEAMEEIVTEAFSISEPGDTVLLAPAAASLDMFKGMGQRGDLFAHNIIGTI
KGLTEEKG"

gene 336. .1808
/gene="ftsW"

CDS 336. .1808
/gene="ftsW"
/function="cell division"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="FtsW protein"
/protein_id="CAB66323.1"
/db_xref="GI:6723457"
/translation="MTTGASKKPARPNTGAKTRTGLGIRERISGAWNDLLARPLTDYI
MILCIVVILSCLGVVMVYSSSMTWSLREGGSVWGTAVRQGIMIVLGFFAMWVALMTRP
QTIRNLSNLILIVSIVLLLAVQIPGIGTGKEEVGSQSWIALGPIQFQPSEIAKVAIAV
WGAHYLAGKGPVQHWFNNHLMREFGGVGAFMAFLIFMEGDAGMAMSFVLVVLFMLFFAG
IAMGWIAIAGVLIIAALAVLAMGGGFRSSRFEVYFDALFGNFHDVRGIAFQSYQGFLS
LADGSGLGVGLGQSRAKWPYLPEAKNDFIFAIIGEELGLWGGALVIALFAGLLYFGLR
TAKKSHDPFLGLMAATLTASVVSQAFINIGYVVGLLPVTGIQLPMISAGGTSAIITLA
SMGLLISCARHEPETVSAMASYGRPAIDRLLGLREPSSTLTTSNASLRSNKTKAAKQK
SESSERVSGPLWRACDRTPSAGATKWASWSTIGSSATLDW"

gene 1994. .3112
/gene="murG"

CDS 1994. .3112
/gene="murG"
/function="peptidoglycan biosynthesis"
/codon_start=1
/transl_table=11
/product="MurG protein"
/protein_id="CAB66324.1"
/db_xref="GI:6723458"
/translation="MKSGELRIKMANSPKPMRVVVAGGGTAGHIEPALAVAEALRDKH
GATVSALGTARGLETTLVPDRGFELHLIEPVPVPRKPNMDLLKLPFRVAKALGQARKA
LKDTDAQAVIGFGGYVSRPAYMAAKSLGLPFFVHEANARAGMANKLGVKLGGVGLNAV
AGSGMDGDVVGIPIRAVLSGARDESAADRARDTWGLDKDRQTIFVTGGSQGSVSINKA
VEQAIDQLVEAGFQVLHAVGKKNELPAAKPGYHPVPFIDDMQAAYTVADLIVCRSGRM
TVAEVTAAGVPAIYVPLPHGNGEQALNAQAVIKAGAARQIDDADFTAQTLIDATLDIL
LHPSTRQSMSDAAKTSTAGNASTVIADMIAATINSQHN"

gene 3147. .4607
/gene="murC"

CDS 3147. .4607
/gene="murC"
/function="peptidoglycan biosynthesis"
/codon_start=1
/transl_table=11
/product="MurC protein"
/protein_id="CAB66325.1"
/db_xref="GI:6723459"
/translation="MTTPHLDSAQDIDLSRVHLIGIGGAGMSGVARILLARGKTVTGS
DAKDSRTLLPLRAVGATIAVGHAAENLELSGELPTVVVTSPAAIPQDNPELVRAREEQ
IPVIRRSDLLGELLEGSTQVLIAGTHGKTSTTSMSVVAMQAAGMDPSFAIGGQLNKAG
TNAHHGTGEVFIAEADESDASLLRYKPNVAVVTNVEPDHLDFFKTPEAYFQVFDDFAG
RITPNGKLVVCLNDPHAAELGERSVRKGIKTVGYGTRDAVQAHPEVPAMATVVDSQVV
AEGTRATINIDGQEVSVILQIPGDHMVLNGAAALLAGYLVGGDVDKLVEGLSDFSGVR
RRFEFHGAIEGGKFNGAAIYDDYAHHPTEVTAVLRRCATRVKAAGKGRVIVAFQPHLY
SRTMEFQKEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVYEPNFSAV
PERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"

BASE COUNT 989 a 1254 c 1317 g 1062 t
ORIGIN

Query Match 60.2%; Score 951; DB 1; Length 4622;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 gcaggtaacgcctccacggtgattgcagacatgattgctgcaactatcaatagccaacac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3047 GCAGGTAACGCCTCCACGGTGATTGCAGACATGATTGCTGCAACTATCAATAGCCAACAC 3106

Qy      61 aactaaaacgaccagctcaacgcaaaggaatagtttaaaggtgaccactccacacttgga 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3107 AACTAAAACGACCAGCTCAACGCAAAGGAATAGTTTAAAGGTGACCACTCCACACTTGGA 3166

Qy     121 ttctgcacaagatattgatctgtcccgcgtccacctcatcggtattggcggagccggaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3167 TTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATCGGTATTGGCGGAGCCGGAAT 3226

Qy     181 gtctggcgttgcccgaatcctgcttgcccgcggtaagacagtcactggttccgatgccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3227 GTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACAGTCACTGGTTCCGATGCCAA 3286

Qy     241 agattcccgcaccttgcttccactccgcgccgtgggagccaccatcgcagtgggacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3287 AGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCCACCATCGCAGTGGGACACGC 3346

Qy     301 tgcggaaaaccttgagctttccggcgaacttcccaccgtcgtggtgacctcttttgccgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3347 TGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTCGTGGTGACCTCTTTTGCCGC 3406

Qy     361 cattccgcaagacaacccggaacttgttcgtgcacgtgaagaaggcattccggttattcg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3407 CATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAAGAAGGCATTCCGGTTATTCG 3466

Qy     421 tcgctccgatctgttgggcgaattgctggaaggctccacccaggtcttgatcgcgggtac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3467 TCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACCCAGGTCTTGATCGCGGGTAC 3526

Qy     481 ccacggtaagacctccaccacctctatgtctgtggtagctatgcaggcagcgggcatgga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3527 CCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCTATGCAGGCAGCGGGCATGGA 3586

Qy     541 tccaagctttgctatcggcggacagctcaacaaggctggcaccaatgcgcaccatggaac 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3587 TCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGCACCAATGCGCACCATGGAAC 3646

Qy     601 tggtgaggtctttatcgctgaagcagatgaatctgacgcatcgctgctgcgctacaagcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3647 TGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCATCGCTGCTGCGCTACAAGCC 3706

Qy     661 aaatgttgcagtggtcaccaatgtggaaccagaccacctggacttctttaaaaccccctga 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3707 AAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTGGACTTCTTTAAAACCCCTGA 3766

Qy     721 agcctacttccaagtgttcgacgatttcgcaggacgcatcaccccgaacggcaagctggt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3767 AGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATCACCCCGAACGGCAAGCTGGT 3826

Qy     781 tgtgtgcctgaacgatcctcacgcagcggagctgggggagaggtctgtccgcaagggtat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3827 TGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAGAGGTCTGTCCGCAAGGGTAT 3886
```

```
Qy    841 caagactgttggttacggtaccgctgacgccgtacaggcacaccctgaggttccagcgat 900
          ||||||||||||||||||||  |||||| ||||||||||||||||||||||||||||||
Db   3887 CAAGACTGTTGGTTACGGTACCCGTGACGCAGTACAGGCACACCCTGAGGTTCCAGCGAT 3946

Qy    901 ggctaccatcgtggattcccaagttgtcgcagaaggcacccgcgccaccatcaacatcga 960
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3947 GGCTACCGTCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCATCAACATCGA 4006

Qy    961 tggacaggaagtatctgtgattcttcaaatccctggtgatcacatggtactcaacggtgc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4007 TGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACTCAACGGTGC 4066

Qy   1021 agccgccctgctggccggatacctggtgggtggggacgtcgacaagcttgttgaaggctt 1080
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   4067 AGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGATAAGCTTGTTGAAGGCTT 4126

Qy   1081 gtcggatttctccggcgtgcgacgccgctttgagttccacggtgctatcgagggcggcaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4127 GTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGAGGGCGGCAA 4186

Qy   1141 atttaatggcgctgctatttatgatgattacgcacaccacccaacggaagtaactgcagt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4187 ATTTAATGGCGCTGCTATTTATGATGATTACGCACACCACCCAACGGAAGTAACTGCAGT 4246

Qy   1201 gctcag 1206
          ||||||
Db   4247 GCTCAG 4252
```

```
RESULT   6
AB003132
LOCUS       AB003132     4116 bp    DNA             BCT       04-AUG-1997
DEFINITION  Corynebacterium glutamicum gene for MurC, FtsQ, FtsZ, complete cds.
ACCESSION   AB003132
VERSION     AB003132.1  GI:2308989
KEYWORDS    FtsZ; FtsQ; MurC.
SOURCE      Corynebacterium glutamicum DNA.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 4116)
  AUTHORS   Kobayashi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1997) to the DDBJ/EMBL/GenBank databases. Miki
            Kobayashi, Mitsubishi Chemical Corp., Tsukuba Research Center; Ami
            8-3-1, Inashiki, Ibaraki, Inashiki, Ibaraki 300-03, Japan
            (E-mail:3709292@cc.m-kagaku.co.jp, Tel:0298-87-1011,
            Fax:0298-87-3259)
REFERENCE   2  (sites)
  AUTHORS   Kobayashi,M., Asai,Y., Hatakeyama,K., Kijima,N., Wachi,M., Nagai,K.
            and Yukawa,H.
  TITLE     Cloning, sequencing, and characterization of the ftsZ gene from
            coryneform bacteria
  JOURNAL   Biochem. Biophys. Res. Commun. 236 (2), 383-388 (1997)
  MEDLINE   97382442
FEATURES             Location/Qualifiers
     source          1. .4116
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            1. .273
                     /gene="murC"
     CDS             <1. .273
                     /gene="murC"
                     /codon_start=1
                     /transl_table=11
                     /product="MurC"
                     /protein_id="BAA21685.1"
                     /db_xref="GI:2308990"
                     /translation="EFQKEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIP
                     VVYEPNFSAVPERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            301. .969
                     /gene="ftsQ"
     CDS             301. .969
                     /gene="ftsQ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsQ"
                     /protein_id="BAA21686.1"
                     /db_xref="GI:2308991"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKAQDAQM
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKRDGQRWNISSPSM
                     VTVR"
     gene            1220. .2536
                     /gene="ftsZ"
     CDS             1220. .2536
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsZ"
                     /protein_id="BAA21687.1"
                     /db_xref="GI:2308992"
                     /translation="MTSPNNYLAKIKVVGVGGGGDNAVNRMIEEGLKGVEFIAVNTDS
                     QALMFSDADVKLDIGREATRGLGAGANPEVGRASAEYHKNEIEETIKGADMVFVTAGE
                     GGGAGTGAAPVVGRIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAALKEVCDTLIV
                     IPNDRLLELGDANLSIMEAFRAADEVLHNGVQGITNLITIPCINVDFADVRSVMSEAG
                     SALMGVGSARGDNRVVSATEQAINSPLLEATMDGATGVLLSFAGGSDLGLMEVNAAAS
                     MVRERSDEDVNLIFGTIIDDNLGDEVRVTVIATGFDAARASAAENRRAGIPAAPAAEP
                     VQQQQVPTTNATLPPEKESIFGGAREENDPYLSRSAGARHRIEETRSGGGLFTTGNDR
                     DYRRDERREDHRDERRDDRSYDRRDDRRDDRRDDRGDDLDVPSFLQ"
     CDS             2564. .3304
                     /note="hypothetical protein similar to YFIH"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA21688.1"
                     /db_xref="GI:2308993"
                     /translation="MDSLDPRNRPVRKVFTTRAGGVSQSPYATFNLGDHVGYYPKAVA
                     SNRNRLADIIGLSPDKVVYMEQIHSNTVTVIDEAPADGQAVEATDALVTTQRGLALAV
                     LVADCVPVLLSDTDAGVIAAVHAGRMGARNGIVAKTIAKMEELGAKPSRIHALMGAAA
                     SGAKYEVPEAMARDVEAKLPGSLARTTKGTTGLDIRAGLLRQMLSLGVQMIDSDPRCT
                     IEDEDLFSYRREGTTGRQAGVVWLPKEA"
     CDS             3345. .4010
                     /note="hypothetical protein similar to YPI1 or YPT5"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA21689.1"
                     /db_xref="GI:2308994"
                     /translation="MQARIDATLNEHNRPEGSVRLLPVTKFHPVEDIKILQELGVTAV
                     GENRPQEARAKALELPDMDFHMIGQIQSKKANSIARWSAAVHSVDSEKIAEALGRGVA
                     LALERGDRTSDELPCFIQLSLDGDPSRGGTPLSQVTQLADCISDPTHLRFEGLMCVPP
                     LGWDPEKAFSQARDVLSGLEEHFDRSLEFSAGMSGDLVAAIKHGSTIVRVGTEILGNR
                     PLA"
BASE COUNT     1009 a    1106 c    1148 g     853 t
ORIGIN
```

QP 501.543

102b

```
  Query Match           14.9%;  Score 236;  DB 1;  Length 4116;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-119;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1346 gagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgat 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58 GAGATTTACGGAGCGCGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGAT 117

Qy   1406 gcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccagaacgcattgca 1465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    118 GCGATGACCATTCCAGTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCA 177

Qy   1466 gaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgtgaccatgctt 1525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    178 GAAATCGCAGGACCTAATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTT 237

Qy   1526 gctccagaaatcctggatcagctgcaaaacaattaggacgtaagtgaacaaggcag 1581
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    238 GCTCCAGAAATCCTGGATCAGCTGCAAAACAATTAGGACGTAAGTGAACAAGGCAG 293


RESULT   7
BLFTSZ
LOCUS       BLFTSZ       5546 bp    DNA             BCT       19-JAN-2000
DEFINITION  Brevibacterium lactofermentum murC(partial), ftsQ, ftsZ genes and
            ORF5, ORF6 DNA.
ACCESSION   Y08964
VERSION     Y08964.2  GI:6723445
KEYWORDS    cell division protein; divB gene; division initiation protein; ftsQ
            gene; ftsZ gene; murC gene; putative YAK 1 protein; putative YPT5
            protein; UDP-N-acetylmuramate-alanine ligase.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 5546)
  AUTHORS   Honrubia,M.P., Fernandez,F.J. and Gil,J.A.
  TITLE     Identification, characterization, and chromosomal organization of
            the ftsZ gene from Brevibacterium lactofermentum
  JOURNAL   Mol. Gen. Genet. 259 (1), 97-104 (1998)
  MEDLINE   98409265
REFERENCE   2  (bases 1 to 5546)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1996) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 5546)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-1999) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
COMMENT     On Jan 20, 2000 this sequence version replaced gi:1769958.
FEATURES             Location/Qualifiers
     source          1. .5546
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            1. .264
                     /gene="murC"
     CDS             <1. .264
                     /gene="murC"
                     /EC_number="6.3.2.8"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-acetylmuramate--alanine ligase"
                     /protein_id="CAA70160.1"
                     /db_xref="GI:1769959"
                     /db_xref="SPTREMBL:P94335"
                     /translation="KEPAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVY
                     EPNFSAVPERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            292. .960
                     /gene="ftsQ"
     CDS             292. .960
                     /gene="ftsQ"
                     /function="division initiation protein or cell division
                     protein"
                     /note="alternative gene name divB"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70161.1"
                     /db_xref="GI:1769960"
                     /db_xref="SPTREMBL:P94336"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKRQDAQM
                     TESIQVVEAPDQPDILLKMNDGREIYWGSSENNHDKAVAMSTVLKREGQRWNISSPSM
                     VTVR"
     gene            1211. .2527
                     /gene="ftsZ"
     CDS             1211. .2527
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein"
                     /protein_id="CAA70158.1"
                     /db_xref="GI:1769961"
                     /db_xref="SWISS-PROT:P94337"
                     /translation="MTSPNNYLAKIKVVGVGGGGVNAVNRMIEEGLKGVEFIAVNTDS
                     QALMFSDADVKLDIGREATRGLGAGANPEVGRASAEDHKNEIEETIKGADMVFVTAGE
                     GGGTGTGAAPVVAGIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAALKEVCDTLIV
                     IPNDRLLELGDANLSIMERFRAADEVLHNGVQGITNLITIPGVINVDFADVRSVMSEA
                     GSALMGVGSARGDNRVVSATEQAINSPLLEATMDGATGVLLSFAGGSDLGLMEVNAAA
                     SMVRERSDEDVNLIFGTIIDDNLGDEVRVTVIATGFDAARASAAENRRAGIPAAPAAE
                     PVQQQVPTTNATLPPEKESIPGGAREENDPYLSRSAGARHRIEETRSGGGLFTTGNDR
                     DYRRDERREDHRDERRDDRSYDRRDDRRDDRRDDRGDDLDVPSFLQ"
     gene            2555. .3295
                     /gene="yfih"
     CDS             2555. .3295
                     /gene="yfih"
                     /note="hypothetical protein similar to YFIH protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70159.1"
                     /db_xref="GI:1769962"
                     /db_xref="SPTREMBL:P94338"
                     /translation="MDSLDPRNRPVRKVPTTRAGGVSQSPYASFNLGDHVGDDPQAVA
                     SNRNRLANIIGLSPDKVVYMEQIHSNTITVIDEAPADGQAVEATDALVTTQRGLALAV
                     LVADCVPVLLSDTDAGVIAAVHAGRMGARNGIVAKTIAKMEELGAKPSRIRALMGAAA
                     SGANYEVPEAMARDVEAKLPGSIARTTKGTTGLDIRAGLLRQMLSLGVQMIDSDPRCT
                     IEDEDLFSYRREGTTGRQAGVVWLPKEA"
     CDS             3376. .4041
                     /note="ORF5"
                     /codon_start=1
                     /transl_table=11
                     /product="putative YPT5 protein"
                     /protein_id="CAB66326.1"
                     /db_xref="GI:6723446"
                     /translation="MQARIDATLNEHNRPEGSVRLLPVTKFHPVEDIKILQELGVTAV
                     GENREQEARAKALELPDMDFHMIGQIQSKKANSIARWAAAVHSVDSEKIAEALGRGVA
                     LALERGDRTGDELPCFIQLSLDGDPSRGGTPLSQVTQLADCISDTTHLRFEGLMCVPP
                     LGWDPEKAFSQARDVLSGLEEHFDRSLEFSAGMSGDLIAAIKHGSTIVRVGTEILGNR
                     PLA"
     CDS             4118. .4576
                     /note="ORF6"
                     /codon_start=1
                     /transl_table=11
                     /product="putative YAK 1 protein"
                     /protein_id="CAB66327.1"
                     /db_xref="GI:6723447"
                     /translation="MSMLKKTKEFFGLAPYEAEHEDAYYADEPRYEGTAAYAPEYRER
                     DYGYAPEAPAPVAPSPAPRSYQSTIVPVELHSFEDAQVIGGAPRDGDAVVFDMSLLSR
                     EEARRIVDFAAGLSFALRGKMQKIDSVTFAVVPELSNISTSELERAARIR"
BASE COUNT     1317 a    1479 c    1520 g    1230 t
ORIGIN


  Query Match           14.9%;  Score 236;  DB 1;  Length 5546;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-119;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1346 gagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgat 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 GAGATTTACGGAGCGCGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGAT 108

Qy   1406 gcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccagaacgcattgca 1465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 GCGATGACCATTCCAGTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCA 168

Qy   1466 gaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgtgaccatgctt 1525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 GAAATCGCAGGACCTAATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTT 228

Qy   1526 gctccagaaatcctggatcagctgcaaaacaattaggacgtaagtgaacaaggcag 1581
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 GCTCCAGAAATCCTGGATCAGCTGCAAAACAATTAGGACGTAAGTGAACAAGGCAG 284


RESULT   8
AX122447
LOCUS       AX122447     1116 bp    DNA             PAT       11-MAY-2001
DEFINITION  Sequence 2363 from Patent EP1108790.
ACCESSION   AX122447
VERSION     AX122447.1  GI:14039694
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 1116)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 2363 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1. .1116
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
BASE COUNT      234 a    317 c    319 g    246 t
ORIGIN


  Query Match             4.0%;  Score 63;  DB 6;  Length 1116;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-23;
  Matches   63;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gcaggtaacgcctccacggtgattgcagacatgattgctgcaactatcaatagccaacac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 GCAGGTAACGCCTCCACGGTGATTGCAGACATGATTGCTGCAACTATCAATAGCCAACAC 1113

Qy     61 aac 63
          |||
Db   1114 AAC 1116


RESULT   9
AF022730
LOCUS       AF022730     1073 bp    mRNA            PLN       30-OCT-1997
DEFINITION  Oryza sativa glyceralehyde-3-phosphate dehydrogenase subunit
            (GAPDH) mRNA, partial cds.
ACCESSION   AF022730
VERSION     AF022730.1  GI:2570494
KEYWORDS    .
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1073)
  AUTHORS   Lee,M.C., Kim,C.S. and Eun,M.Y.
  TITLE     Isolation and characterization of Glyceraldehyde-3-phosphate
            dehydrogenase from rice
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1073)
  AUTHORS   Lee,M.C., Kim,C.S. and Eun,M.Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-SEP-1997) Cytogenetics, National Institute of
            Agricultural Science and Technology, RDA, Sedun-Dong, Suwon
            441-707, Korea
FEATURES             Location/Qualifiers
     source          1. .1073
                     /organism="Oryza sativa"
                     /strain="Milyang 23"
                     /db_xref="taxon:4530"
                     /tissue_type="seed"
                     /dev_stage="immature (milky stage)"
                     /clone="KCDM2189"
     gene            1. .>1073
                     /gene="GAPDH"
     CDS             45. .>1073
                     /gene="GAPDH"
                     /codon_start=1
                     /product="glyceralehyde-3-phosphate dehydrogenase subunit"
                     /protein_id="AAB82133.1"
                     /db_xref="GI:2570495"
                     /translation="MASPCSPPPPCHSRGGGLSDPSGLRSSSSLPLRRDATSDNFMNP
                     VSFRTHAVGTNGGARRAPTEAKLKVAINGFGRFGRNFLRCWHGRGDDSPLNIIPINDT
                     GGVKQASHRLKYNSTLGIFDADVKPVGENPLSVEGEVIKVVSDRNPSNLPWGELGIEL
                     LIEGTGLFFNANGAGKHIQAGAKKVLITAPGKGDIPTYVVGVNADQYSPDEPIISNAS
                     CTTNCLAPFVKILDQKFGIIKGTMTTTHSYTGDQRVFDGSHRDLRRARAAALNIVPTS
                     TGAAKAVALVLPNLKGKLNGIALRVPTPNVSVVDLVVQVSKKTLAEEVNQAFRDSAAN
                     ELKASSRER"
BASE COUNT      196 a    380 c    335 g    162 t
ORIGIN


  Query Match             1.3%;  Score 21;  DB 8;  Length 1073;
  Best Local Similarity   100.0%;  Pred. No. 9.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1076 ggcttgtcggatttctccggc 1096
          |||||||||||||||||||||
Db     93 GGCTTGTCGGATTTCTCCGGC 113


RESULT  10
AC019805/c
LOCUS       AC019805     6843 bp    DNA             HTG       03-JAN-2000
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC019805
VERSION     AC019805.1  GI:6665092
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 6843)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10211017 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .6843
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT     1966 a   1466 c   1449 g   1962 t
ORIGIN


  Query Match             1.3%;  Score 21;  DB 2;  Length 6843;
  Best Local Similarity   100.0%;  Pred. No. 7.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    262 actccgcgccgtgggagccac 282
          |||||||||||||||||||||
Db   5976 ACTCCGCGCCGTGGGAGCCAC 5956


RESULT  11
AC010010
```

```
LOCUS       AC010010    85875 bp    DNA             HTG       18-FEB-2000
DEFINITION  Drosophila melanogaster clone RPCI98-13F6, *** SEQUENCING IN
            PROGRESS ***, 53 unordered pieces.
ACCESSION   AC010010
VERSION     AC010010.3  GI:6996694
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 85875)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Farraguto,D.,
            Forcum-Tansey,J., Frantz,P.,. Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
            Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
            Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
            Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
            Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 85875)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Feb 18, 2000 this sequence version replaced gi:5881484.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: DREG
                Center clone name: RPCI98-13F6
            ---------- Summary Statistics
                Sequencing vector: M13; L08821
                Chemistry: Dye-primer Bodipy: 68% of reads
                Chemistry: Dye-terminator Big Dye: 31% of reads
                Assembly program: Phrap; version 0.980611
                Consensus quality: 41466 bases at least Q40
                Consensus quality: 57245 bases at least Q30
                Consensus quality: 64739 bases at least Q20
                Estimated insert size: 75108; sum-of-contigs estimation
                Quality coverage: 0.8x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 53 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1256: contig of 1256 bp in length
            *     1257     1276: gap of unknown length
            *     1277     2355: contig of 1079 bp in length
            *     2356     2375: gap of unknown length
            *     2376     3452: contig of 1077 bp in length
            *     3453     3472: gap of unknown length
            *     3473     4275: contig of 803 bp in length
            *     4276     4295: gap of unknown length
            *     4296     5562: contig of 1267 bp in length
            *     5563     5582: gap of unknown length
            *     5583     6777: contig of 1195 bp in length
            *     6778     6797: gap of unknown length
            *     6798     8061: contig of 1264 bp in length
            *     8062     8081: gap of unknown length
            *     8082     9053: contig of 972 bp in length
            *     9054     9073: gap of unknown length
            *     9074     9934: contig of 861 bp in length
            *     9935     9954: gap of unknown length
            *     9955    10766: contig of 812 bp in length
            *    10767    10786: gap of unknown length
            *    10787    11616: contig of 830 bp in length
            *    11617    11636: gap of unknown length
            *    11637    12464: contig of 828 bp in length
            *    12465    12484: gap of unknown length
            *    12485    13645: contig of 1161 bp in length
            *    13646    13665: gap of unknown length
            *    13666    14489: contig of 824 bp in length
            *    14490    14509: gap of unknown length
            *    14510    15861: contig of 1352 bp in length
            *    15862    15881: gap of unknown length
            *    15882    16921: contig of 1040 bp in length
            *    16922    16941: gap of unknown length
            *    16942    18465: contig of 1524 bp in length
            *    18466    18485: gap of unknown length
            *    18486    19316: contig of 831 bp in length
            *    19317    19336: gap of unknown length
            *    19337    20903: contig of 1567 bp in length
            *    20904    20923: gap of unknown length
            *    20924    22150: contig of 1227 bp in length
            *    22151    22170: gap of unknown length
            *    22171    23295: contig of 1125 bp in length
            *    23296    23315: gap of unknown length
            *    23316    24310: contig of 995 bp in length
            *    24311    24330: gap of unknown length
            *    24331    26217: contig of 1887 bp in length
            *    26218    26237: gap of unknown length
            *    26238    27815: contig of 1578 bp in length
            *    27816    27835: gap of unknown length
            *    27836    28794: contig of 959 bp in length
            *    28795    28814: gap of unknown length
            *    28815    30095: contig of 1281 bp in length
            *    30096    30115: gap of unknown length
            *    30116    31085: contig of 970 bp in length
            *    31086    31105: gap of unknown length
            *    31106    32390: contig of 1285 bp in length
            *    32391    32410: gap of unknown length
            *    32411    34245: contig of 1835 bp in length
            *    34246    34265: gap of unknown length
            *    34266    35631: contig of 1366 bp in length
            *    35632    35651: gap of unknown length
            *    35652    37059: contig of 1408 bp in length
            *    37060    37079: gap of unknown length
            *    37080    38607: contig of 1528 bp in length
            *    38608    38627: gap of unknown length
            *    38628    40420: contig of 1793 bp in length
            *    40421    40440: gap of unknown length
            *    40441    41701: contig of 1261 bp in length
            *    41702    41721: gap of unknown length
            *    41722    43319: contig of 1598 bp in length
            *    43320    43339: gap of unknown length
            *    43340    45717: contig of 2378 bp in length
            *    45718    45737: gap of unknown length
            *    45738    47273: contig of 1536 bp in length
            *    47274    47293: gap of unknown length
            *    47294    48892: contig of 1599 bp in length
            *    48893    48912: gap of unknown length
            *    48913    50344: contig of 1432 bp in length
            *    50345    50364: gap of unknown length
            *    50365    52143: contig of 1779 bp in length
```

```
            *    52144    52163: gap of unknown length
            *    52164    54492: contig of 2329 bp in length
            *    54493    54512: gap of unknown length
            *    54513    56276: contig of 1764 bp in length
            *    56277    56296: gap of unknown length
            *    56297    57664: contig of 1368 bp in length
            *    57665    57684: gap of unknown length
            *    57685    59691: contig of 2007 bp in length
            *    59692    59711: gap of unknown length
            *    59712    63581: contig of 3870 bp in length
            *    63582    63601: gap of unknown length
            *    63602    65422: contig of 1821 bp in length
            *    65423    65442: gap of unknown length
            *    65443    67456: contig of 2014 bp in length
            *    67457    67476: gap of unknown length
            *    67477    69545: contig of 2069 bp in length
            *    69546    69565: gap of unknown length
            *    69566    72562: contig of 2997 bp in length
            *    72563    72582: gap of unknown length
            *    72583    75310: contig of 2728 bp in length
            *    75311    75330: gap of unknown length
            *    75331    77972: contig of 2642 bp in length
            *    77973    77992: gap of unknown length
            *    77993    82474: contig of 4482 bp in length
            *    82475    82494: gap of unknown length
            *    82495    85875: contig of 3381 bp in length.
FEATURES             Location/Qualifiers
     source          1. .85875
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RPCI98-13F6"
BASE COUNT    23385 a  18706 c  18751 g  23920 t   1113 others
ORIGIN


  Query Match             1.3%;  Score 21;  DB 2;  Length 85875;
  Best Local Similarity   100.0%;  Pred. No. 6.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    262 actccgcgccgtgggagccac 282
          |||||||||||||||||||||
Db  37850 ACTCCGCGCCGTGGGAGCCAC 37870


RESULT  12
AC012665/c
LOCUS       AC012665   126039 bp    DNA             PRI       03-JUL-2001
DEFINITION  Homo sapiens clone RP11-19E11, complete sequence.
ACCESSION   AC012665
VERSION     AC012665.8  GI:14589661
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 126039)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 126039)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 126039)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 3, 2001 this sequence version replaced gi:13794251.
            ------------------ Genome Center ------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            --------------- Project Information ----------------
            Center project name: H_NH0019E11
            ----------------------------------------------------
FEATURES             Location/Qualifiers
     source          1. .126039
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-19E11"
BASE COUNT    33691 a  29032 c  28371 g  34945 t
ORIGIN


  Query Match             1.3%;  Score 21;  DB 9;  Length 126039;
  Best Local Similarity   100.0%;  Pred. No. 6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1537 cctggatcagctgcaaaacaa 1557
          |||||||||||||||||||||
Db 115904 CCTGGATCAGCTGCAAAACAA 115884


RESULT  13
AC084093/c
LOCUS       AC084093   137569 bp    DNA             HTG       09-MAY-2001
DEFINITION  Homo sapiens chromosome 8 clone CTD-2277K12 map 8, WORKING DRAFT
            SEQUENCE, 3 unordered pieces.
ACCESSION   AC084093
VERSION     AC084093.2  GI:13123904
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 137569)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone CTD-2277K12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 137569)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
            Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
            Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,
            Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihóva,T., Mlenga,V.,
            Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
            Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 26, 2001 this sequence version replaced gi:10799458.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
```

```
            Contact: sequence_submissions@genome.wi.mit.edu
            ---------- Project Information
            Center project name: L11391
            Center clone name: 2277_K_12
            ---------- Summary Statistics
            Sequencing vector: Plasmid; n/a; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 136529 bases at least Q40
            Consensus quality: 137130 bases at least Q30
            Consensus quality: 137294 bases at least Q20
            Insert size: 139000; agarose-fp
            Insert size: 137369; sum-of-contigs
            Quality coverage: 7.8 in Q20 bases; agarose-fp
            Quality coverage: 7.9 in Q20 bases; sum-of-contigs
            ----------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    37849: contig of 37849 bp in length
            *   37850 37949: gap of      100 bp
            *   37950    82422: contig of 44473 bp in length
            *   82423 82522: gap of      100 bp
            *   82523   137569: contig of 55047 bp in length.
FEATURES             Location/Qualifiers
     source          1. .137569
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="CTD-2277K12"
                     /clone_lib="CITD Human BAC"
     misc_feature    1. .37849
                     /note="assembly_fragment"
     misc_feature    37950. .82422
                     /note="assembly_fragment"
     misc_feature    82523. .137569
                     /note="assembly_fragment"
BASE COUNT    42095 a  26887 c  26629 g  41757 t    201 others
ORIGIN


  Query Match             1.3%;  Score 21;  DB 2;  Length 137569;
  Best Local Similarity   100.0%;  Pred. No. 6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     863 gctgacgccgtacaggcacac 883
           |||||||||||||||||||||
Db   33227 GCTGACGCCGTACAGGCACAC 33207


RESULT  14
AC022846
LOCUS       AC022846    163027 bp    DNA             HTG       09-MAY-2001
DEFINITION  Homo sapiens chromosome 8 clone RP11-253N21 map 8, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC022846
VERSION     AC022846.3  GI:13123397
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163027)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-253N21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163027)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 25, 2001 this sequence version replaced gi:7139770.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ---------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            ---------- Project Information
                Center project name: L6163
                Center clone name: 253_N_21
            ---------- Summary Statistics
                Sequencing vector: M13; M77815; 39% of reads
                Sequencing vector: Plasmid; n/a; 61% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 160669 bases at least Q40
                Consensus quality: 161671 bases at least Q30
                Consensus quality: 162136 bases at least Q20
                Insert size: 157000; agarose-fp
                Insert size: 162527; sum-of-contigs
                Quality coverage: 13.6 in Q20 bases; agarose-fp
                Quality coverage: 13.1 in Q20.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    23967: contig of 23967 bp in length
            *   23968 24067: gap of      100 bp
            *   24068    25128: contig of 1061 bp in length
            *   25129 25228: gap of      100 bp
            *   25229    34037: contig of 8809 bp in length
            *   34038 34137: gap of      100 bp
            *   34138    56813: contig of 22676 bp in length
            *   56814 56913: gap of      100 bp
            *   56914    92224: contig of 35311 bp in length
            *   92225 92324: gap of      100 bp
            *   92325   163027: contig of 70703 bp in length.
FEATURES             Location/Qualifiers
     source          1. .163027
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-253N21"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .23967
                     /note="assembly_fragment
```

```
                     clone_end:SP6
                     vector_side:left"
     misc_feature    24068. .25128
                     /note="assembly_fragment"
     misc_feature    25229. .34037
                     /note="assembly_fragment"
     misc_feature    34138. .56813
                     /note="assembly_fragment"
     misc_feature    56914. .92224
                     /note="assembly_fragment"
     misc_feature    92325. .163027
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
BASE COUNT    48042 a  32608 c  32206 g  49671 t    500 others
ORIGIN


  Query Match             1.3%;  Score 21;  DB 2;  Length 163027;
  Best Local Similarity   100.0%;  Pred. No. 5.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    863 gctgacgccgtacaggcacac 883
          |||||||||||||||||||||
Db  90055 GCTGACGCCGTACAGGCACAC 90075


RESULT  15
AC013275
LOCUS       AC013275   174822 bp    DNA             PRI       09-AUG-2001
DEFINITION  Homo sapiens BAC clone RP11-464C8 from 2, complete sequence.
ACCESSION   AC013275
VERSION     AC013275.8  GI:15144471
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 174822)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome research. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 174822)
  AUTHORS   Sun,H., Haakenson,W. and Conners,J.
  TITLE     The sequence of Homo sapiens BAC clone RP11-464C8
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 174822)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 174822)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Aug 9, 2001 this sequence version replaced gi:13270787.
            ---------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            ---------- Summary Statistics
                Center project name: H_NH0464C08
            ----------

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-393J17; the clone sequenced
            to the right is AC016076.  Actual start of this clone is at base
            position 1 of RP11-464C8; actual end is at base position 174822 of
            RP11-464C8.

            The region between 76913 to 76922 is covered only by a single m13
            subclone.  Digest information confirms the assembly.
FEATURES             Location/Qualifiers
     source          1. .174822
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-464C8"
                     /clone_lib="RPCI-11"
     repeat_region   1305. .1548
                     /rpt_family="MER2_type"
     repeat_region   1864. .1954
                     /rpt_family="MIR"
     repeat_region   1955. .2254
                     /rpt_family="Alu"
     repeat_region   2255. .2334
                     /rpt_family="MIR"
     misc_feature    3496. .4429
                     /note="similar to  EST AL589988 (NID:g13396825)"
     repeat_region   5380. .5666
                     /rpt_family="Alu"
     repeat_region   6188. .6501
                     /rpt_family="Alu"
     repeat_region   6975. .6998
                     /rpt_family="(TTTG)n"
     repeat_region   8647. .8809
                     /rpt_family="Alu"
     repeat_region   8823. .8884
                     /rpt_family="MER2_type"
     repeat_region   9261. .9396
                     /rpt_family="MER1_type"
     repeat_region   9397. .9711
                     /rpt_family="Alu"
     repeat_region   9712. .9758
                     /rpt_family="MER1_type"
     repeat_region   10174. .10321
                     /rpt_family="Alu"
     repeat_region   10322. .10366
                     /rpt_family="(CA)n"
     repeat_region   10367. .10453
                     /rpt_family="Alu"
     repeat_region   10467. .10527
```

```
                     /rpt_family="MER1_type"
repeat_region        10604. .10930
                     /rpt_family="ERV1"
repeat_region        10931. .10942
                     /rpt_family="Alu"
repeat_region        10943. .10974
                     /rpt_family="(TG)n"
repeat_region        10975. .11271
                     /rpt_family="Alu"
repeat_region        11272. .11629
                     /rpt_family="ERV1"
repeat_region        11659. .11925
                     /rpt_family="ERV1"
repeat_region        12011. .12316
                     /rpt_family="Alu"
repeat_region        12389. .12763
                     /rpt_family="ERV1"
repeat_region        12764. .12798
                     /rpt_family="MaLR"
repeat_region        12799. .13097
                     /rpt_family="Alu"
repeat_region        13098. .13493
                     /rpt_family="MaLR"
repeat_region        13494. .13832
                     /rpt_family="MaLR"
repeat_region        13833. .14250
                     /rpt_family="ERV1"
repeat_region        14251. .15502
                     /rpt_family="MaLR"
repeat_region        15505. .15912
                     /rpt_family="MaLR"
repeat_region        15968. .16092
                     /rpt_family="ERV1"
repeat_region        16797. .17059
                     /rpt_family="MER1_type"
repeat_region        18122. .18307
                     /rpt_family="MER1_type"
repeat_region        18309. .18885
                     /rpt_family="ERV1"
repeat_region        18953. .19126
                     /rpt_family="Alu"
misc_feature         19909. .20210
                     /note="similar to  EST AA668276 (NID:g2629775) ab78e11.s1"
repeat_region        20844. .20938
                     /rpt_family="MIR"
repeat_region        20988. .21202
                     /rpt_family="L2"
repeat_region        21545. .21643
                     /rpt_family="L2"
repeat_region        21625. .21968
                     /rpt_family="L2"
repeat_region        22407. .22662
                     /rpt_family="ERV1"
repeat_region        22689. .22871
                     /rpt_family="ERV1"
repeat_region        23543. .23662
                     /rpt_family="MER1_type"
repeat_region        23663. .23833
                     /rpt_family="MIR"
repeat_region        24159. .24465
                     /rpt_family="Alu"
misc_feature         24509. .24802
                     /note="similar to  EST AI631785 (NID:g4683115) wa36a10.x1"
misc_feature         24515. .24885
                     /note="similar to  EST BF589995 (NID:g11682319)"
misc_feature         24684. .25493
                     /note="CpG_island (%GC=73.5, o/e=0.80, #CpGs=89)"
misc_feature         25098. .25228
                     /note="similar to  EST BF532525 (NID:g11619888)"
misc_feature         25180. .25228
                     /note="similar to  EST BF787117 (NID:g12092153)"
misc_feature         25180. .25228
                     /note="similar to  EST AI507063 (NID:g4404914) vj32h01.x1"
misc_feature         25180. .25228
                     /note="similar to  EST AW742372 (NID:g7654159) up56c03.y1"
misc_feature         25190. .25642
                     /note="similar to  EST AA771837 (NID:g2823620) a136d03.s1"
repeat_region        26043. .26175
                     /rpt_family="MIR"
repeat_region        26345. .26419
                     /rpt_family="L2"
repeat_region        27624. .27751
                     /rpt_family="MIR"

  Query Match             1.3%;  Score 21;  DB 9;  Length 174822;
  Best Local Similarity   100.0%;  Pred. No. 5.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1537 cctggatcagctgcaaaacaa 1557
          |||||||||||||||||||||
Db   6112 CCTGGATCAGCTGCAAAACAA 6132

Search completed: February 15, 2002, 10:47:55
Job time: 2984 sec
```

```
Db     238 GCTCCAGAAATCCTGGATCAGCTGCAAAACAATTAGGACGTAAGTGAACAAGGCAG 293

RESULT   7
BLFTSZ
LOCUS       BLFTSZ       5546 bp    DNA             BCT       19-JAN-2000
DEFINITION  Brevibacterium lactofermentum murC(partial), ftsQ, ftsZ genes and
            ORF5, ORF6 DNA.
ACCESSION   Y08964
VERSION     Y08964.2  GI:6723445
KEYWORDS    cell division protein; divB gene; division initiation protein; ftsQ
            gene; ftsZ gene; murC gene; putative YAK 1 protein; putative YPT5
            protein; UDP-N-acetylmuramate-alanine ligase.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 5546)
  AUTHORS   Honrubia,M.P., Fernandez,F.J. and Gil,J.A.
  TITLE     Identification, characterization, and chromosomal organization of
            the ftsZ gene from Brevibacterium lactofermentum
  JOURNAL   Mol. Gen. Genet. 259 (1), 97-104 (1998)
  MEDLINE   98409265
REFERENCE   2  (bases 1 to 5546)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1996) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 5546)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-1999) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
COMMENT     On Jan 20, 2000 this sequence version replaced gi:1769958.
FEATURES             Location/Qualifiers
     source          1. .5546
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            1. .264
                     /gene="murC"
     CDS             <1. .264
                     /gene="murC"
                     /EC_number="6.3.2.8"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-acetylmuramate--alanine ligase"
                     /protein_id="CAA70160.1"
                     /db_xref="GI:1769959"
                     /db_xref="SPTREMBL:P94335"
                     /translation="REPAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVY
                     EPNFSAVPERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            292. .960
                     /gene="ftsQ"
     CDS             292. .960
                     /gene="ftsQ"
                     /function="division initiation protein or cell division
                     protein"
                     /note="alternative gene name divB"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70161.1"
                     /db_xref="GI:1769960"
                     /db_xref="SPTREMBL:P94336"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIRRQDAQM
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKREGQRWNISSPSM
                     VTVR"
                     1211. .2527
                     /gene="ftsZ"
                     1211. .2527
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein"
                     /protein_id="CAA70158.1"
                     /db_xref="GI:1769961"
                     /db_xref="SWISS-PROT:P94337"
                     /translation="MTSPNNYLAKIKVVGVGGGGVNAVNRMIEEGLKGVEFIAVNTDS
                     QALMFSDADVKLDIGREATRGLGAGANPEVGRASAEDHKNEIEETIKGADMVFVTAGE
                     GGGTGTGAAPVVAGIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAALKEVCDTLIV
                     IPNDRLLELGDANLSIMERFRAADEVLHNGVQGITNLITIPGVINVDFADVRSVMSEA
                     GSALMGVGSARGDNRVVSATEQAINSPLLEATMDGATGVLLSFAGGSDLGLMEVNAAA
                     SMVRERSDEDVNLIFGTIIDDNLGDEVRVTVIATGFDAARASAAENRRAGIPAAPAAE
                     PVQQQVPTTNATLPPEKESIFGGAREENDPYLSRSAGARHRIEETRSGGGLFTTGNDR
                     DYRRDERREDHRDERRDDRSYDRRDDRRDDRRDDRGDDLDVPSFLQ"
     gene            2555. .3295
                     /gene="yfih"
     CDS             2555. .3295
                     /gene="yfih"
                     /note="hypothetical protein similar to YFIH protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70159.1"
                     /db_xref="GI:1769962"
                     /db_xref="SPTREMBL:P94338"
                     /translation="MDSLDPRNRPVRKVPTTRAGGVSQSPYASFNLGDHVGDDPQAVA
                     SNRNRLANIIGLSPDKVVTMEQIHSNTITVIDEAPADGQAVEATDALVTTQRGLALAV
                     LVADCVPVLLSDTDAGVIAAVHAGRMGARNGIVAKTIAKMEELGAKPSRIHALMGAAA
                     SGANYEVPEAMARDVEAKLPGSIARTTKGTTGLDIRAGLLRQMLSLGVQMIDSDPRCT
                     IEDEDLFSYRREGTTGRQAGVVWLPKEA"
     CDS             3376. .4041
                     /note="ORF5"
                     /codon_start=1
                     /transl_table=11
                     /product="putative YPT5 protein"
                     /protein_id="CAB66326.1"
                     /db_xref="GI:6723446"
                     /translation="MQARIDATLNEHNRPEGSVRLLPVTKFHPVEDIKILQELGVTAV
                     GENREQEARAKALELPDMDFHMIGQIQSKKANSIARWAAAVHSVDSEKIAEALGRQVA
                     LALERGDRTGDELPCFIQLSLDGDPSRGGTPLSQVTQLADCISDTTHLRFEGLMCVPP
                     LGWDPEKAPSQARDVLSGLEEHFDRSLEFSAGMSGDLIAAIKHGSTIVRVGTEILGNR
                     PLA"
     CDS             4118. .4576
                     /note="ORF6"
                     /codon_start=1
                     /transl_table=11
                     /product="putative YAK 1 protein"
                     /protein_id="CAB66327.1"
                     /db_xref="GI:6723447"
                     /translation="MSMLKKTKEPFGLAPYEAEHEDAYYADEPRYEGTAAYAPEYRER
                     DYGYAPEAPAPVAPSPAPRSYQSTIVPVELHSPEDAQVIGGAFRDGDAVVFDMSLLSR
                     EEARRIVDFAAGLSFALRGKMQKIDSVTFAVVPELSNISTSELERAARIR"
BASE COUNT     1317 a    1479 c    1520 g    1230 t
ORIGIN


  Query Match          14.9%;  Score 236;  DB 1;  Length 5546;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-119;
  Matches  236;  Conservative    0;  Mismatches   .0;  Indels    0;  Gaps    0;

Qy   1346 gagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgat 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 GAGATTTACGGAGCGCGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGAT 108

Qy   1406 gcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccagaacgcattgca 1465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 GCGATGACCATTCCAGTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCA 168

Qy   1466 gaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgtgaccatgctt 1525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 GAAATCGCAGGACCTAATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTT 228

Qy   1526 gctccagaaatcctggatcagctgcaaaacaattaggacgtaagtgaacaaggcag 1581
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 10:15:55 ; Search time 49.36 Seconds
(without alignments)
7254.087 Million cell updates/sec

Title: US-09-603-124B-1
Perfect score: 1581
Sequence: 1 gcaggtaacgcctccacggt.....gacgtaagtgaacaaggcag 1581

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	18	1.1	2289	2	US-08-463-081B-30	Sequence 30, Appl
c 2	18	1.1	2289	2	US-08-461-379A-30	Sequence 30, Appl
c 3	18	1.1	2289	2	US-08-462-390B-30	Sequence 30, Appl
c 4	18	1.1	2289	3	US-08-463-074B-30	Sequence 30, Appl
c 5	18	1.1	2289	3	US-08-465-585C-30	Sequence 30, Appl
c 6	18	1.1	2289	3	US-08-652-446-30	Sequence 30, Appl
c 7	18	1.1	2946	1	US-08-473-122-1	Sequence 1, Appli
c 8	18	1.1	2946	2	US-08-472-478-1	Sequence 1, Appli
c 9	18	1.1	2946	2	US-08-463-081B-7	Sequence 7, Appli
c 10	18	1.1	2946	2	US-08-461-379A-7	Sequence 7, Appli
c 11	18	1.1	2946	2	US-08-462-390B-7	Sequence 7, Appli
c 12	18	1.1	2946	3	US-08-463-074B-7	Sequence 7, Appli
c 13	18	1.1	2946	3	US-08-465-585C-7	Sequence 7, Appli
c 14	18	1.1	2946	3	US-08-652-446-7	Sequence 7, Appli
c 15	17	1.1	300	2	US-08-308-494A-16	Sequence 16, Appl
c 16	17	1.1	309	1	US-08-467-393-3	Sequence 3, Appli
c 17	17	1.1	322	3	US-08-434-000A-11	Sequence 11, Appl
c 18	17	1.1	322	4	US-09-312-157-11	Sequence 11, Appl
c 19	17	1.1	324	3	US-08-483-749A-7	Sequence 7, Appli
c 20	17	1.1	327	2	US-08-002-324-3	Sequence 3, Appli
c 21	17	1.1	327	5	PCT-US94-00261-3	Sequence 3, Appli
c 22	17	1.1	331	3	US-08-836-561-32	Sequence 32, Appl
c 23	17	1.1	357	4	US-09-171-945-8	Sequence 8, Appli
c 24	17	1.1	360	1	US-08-447-422-13	Sequence 13, Appl
c 25	17	1.1	384	2	US-08-656-586-1	Sequence 1, Appli
c 26	17	1.1	384	4	US-08-619-491-1	Sequence 1, Appli
c 27	17	1.1	384	5	PCT-US95-07302-1	Sequence 1, Appli
c 28	17	1.1	393	2	US-08-116-778E-39	Sequence 39, Appl
c 29	17	1.1	393	2	US-08-438-562-39	Sequence 39, Appl
c 30	17	1.1	393	2	US-08-483-528B-2	Sequence 2, Appli
c 31	17	1.1	393	3	US-08-673-799C-2	Sequence 2, Appli
c 32	17	1.1	462	1	US-07-946-421-25	Sequence 25, Appl
c 33	17	1.1	646	2	US-08-737-129A-3	Sequence 3, Appli
c 34	17	1.1	648	6	5455030-4	Patent No. 5455030
c 35	17	1.1	669	2	US-08-190-199A-66	Sequence 66, Appl
c 36	17	1.1	672	2	US-08-190-199A-62	Sequence 62, Appl
c 37	17	1.1	705	4	US-09-171-945-16	Sequence 16, Appl
c 38	17	1.1	708	2	US-08-190-199A-60	Sequence 60, Appl
c 39	17	1.1	711	2	US-08-190-199A-64	Sequence 64, Appl
c 40	17	1.1	719	3	US-08-279-772A-7	Sequence 7, Appli
c 41	17	1.1	720	1	US-08-061-092A-2	Sequence 2, Appli
c 42	17	1.1	720	3	US-08-902-486-10	Sequence 10, Appl
c 43	17	1.1	720	6	5455030-14	Patent No. 5455030
c 44	17	1.1	729	6	5455030-16	Patent No. 5455030
c 45	17	1.1	732	2	US-08-553-497A-19	Sequence 19, Appl

ALIGNMENTS

```
RESULT 1
US-08-463-081B-30/c
; Sequence 30, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960 5837487
; GENERAL INFORMATION:
;   APPLICANT: Smith, Kendall A. & Beadling, Carol
;   TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
;   TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression The
;   NUMBER OF SEQUENCES: 35
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
;     STREET: 444 South Flower St. - Suite 1900
;     CITY: Los Angeles
;     STATE: California
;     COUNTRY: USA
;     ZIP: 90071
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0,
;     SOFTWARE: Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/463,081B
;     FILING DATE: 5-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/104,736
;     FILING DATE: 10-AUG-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/796,066
;     FILING DATE: 20-NOV-91
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Viviana Amzel, Ph. D.
;     REGISTRATION NUMBER: 30,930
;     REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (213) 622-7700
;     TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 30:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2289 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-463-081B-30


  Query Match 1.1%; Score 18; DB 2; Length 2289;
```

```
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2113 TGCCGTGGTGCTTGAGAT 2096


RESULT   2
US-08-461-379A-30/c
; Sequence 30, Application US/08461379A
; Patent No. 5871961
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5 Polypeptide,
;    TITLE OF INVENTION:  Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:                (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,379A
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108; 08/104,736
;      APPLICATION NUMBER:  & 07/796,066
;      FILING DATE:  27-OCT-1994; 10-AUG-1993 & 20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-070
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)470-0700
;      TELEFAX:  (610)470-0701
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2289 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-461-379A-30


  Query Match             1.1%;  Score 18;  DB 2;  Length 2289;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2113 TGCCGTGGTGCTTGAGAT 2096


RESULT   3
US-08-462-390B-30/c
; Sequence 30, Application US/08462390B
; Patent No. 5882894
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR8 Polypeptide, Vector and
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:                (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,390B
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)407-0700
;      TELEFAX:  (610)407-0701
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2289 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-462-390B-30


  Query Match             1.1%;  Score 18;  DB 2;  Length 2289;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2113 TGCCGTGGTGCTTGAGAT 2096


RESULT   4
US-08-463-074B-30/c
; Sequence 30, Application US/08463074B
; Patent No. 6020155
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR1 Fusion Protein, Vector an
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:                (B) STREET:                444 South Flower St. - Suit
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,074B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38143 (DART-020)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2289 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-463-074B-30


  Query Match           1.1%;  Score 18;  DB 3;  Length 2289;
  Best Local Similarity  100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2113 TGCCGTGGTGCTTGAGAT 2096


RESULT   5
US-08-465-585C-30/c
; Sequence 30, Application US/08465585C
; Patent No. 6027914
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR6 Polypeptide, Vecto
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY,SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:              444South Flower St. - Suite 190
;      CITY:  Los Angeles
;      STATE:  Californiaa
;      COUNTRY:  USA
;      ZIP:  900071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,585C
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38149 (DART-050)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 4894210
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2289 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-465-585C-30


  Query Match           1.1%;  Score 18;  DB 3;  Length 2289;
  Best Local Similarity  100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2113 TGCCGTGGTGCTTGAGAT 2096


RESULT   6
US-08-652-446-30/c
; Sequence 30, Application US/08652446
; Patent No. 6057427
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5
;    TITLE OF INVENTION:  Polypeptide, Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:              444 South Flower St. - Sui
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,446
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP App. # 96921319.8
;      FILING DATE:  5-JAN-1998
;      APPLICATION NUMBER:  PCT/US/96/09194
;      FILING DATE:  5-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/330,108
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/463,074
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/462,337
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/462,390
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,585
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/463,081
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/461,379
;      FILING DATE:  5-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/739,523
;      FILING DATE:  29-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
```

; REFERENCE/DOCKET NUMBER: FP66 40035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-652-446-30

Query Match 1.1%; Score 18; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1333 tgccgtggtgcttgagat 1350
||||||||||||||||||
Db 2113 TGCCGTGGTGCTTGAGAT 2096

RESULT 7
US-08-473-122-1/c
; Sequence 1, Application US/08473122
; Patent No. 5652340
; GENERAL INFORMATION:
; APPLICANT: KOHWI-SHIGEMATSU, TERUMI
; APPLICANT: KOHWI, YOSHINORI
; APPLICANT: DICKINSON, LILIANE A
; TITLE OF INVENTION: MATRIX-ASSOCIATING DNA-BINDING PROTEIN,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME AND METHODS FOR DETECTING
; TITLE OF INVENTION: THE NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,122
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/934,034
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619 535-9001
; TELEFAX: 619 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-473-122-1

Query Match 1.1%; Score 18; DB 1; Length 2946;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1333 tgccgtggtgcttgagat 1350
||||||||||||||||||
Db 2327 TGCCGTGGTGCTTGAGAT 2310

RESULT 8
US-08-472-478-1/c
; Sequence 1, Application US/08472478
; Patent No. 5869621
; GENERAL INFORMATION:
; APPLICANT: KOHWI-SHIGEMATSU, TERUMI
; APPLICANT: KOHWI, YOSHINORI
; APPLICANT: DICKINSON, LILIANE A
; TITLE OF INVENTION: MATRIX-ASSOCIATING DNA-BINDING PROTEIN,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME AND METHODS FOR DETECTIN
; TITLE OF INVENTION: THE NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,478
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/934,034
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 9293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619 535-9001
; TELEFAX: 619 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2946 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-472-478-1

Query Match 1.1%; Score 18; DB 2; Length 2946;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1333 tgccgtggtgcttgagat 1350
||||||||||||||||||
Db 2327 TGCCGTGGTGCTTGAGAT 2310

RESULT 9
US-08-463-081B-7/c
; Sequence 7, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960 5837487
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling; Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression There

```
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      STREET:  444 South Flower St. - Suite 1900
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,081B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38150 (DART-060)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2946 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  215..2503
US-08-463-081B-7


  Query Match             1.1%;  Score 18;  DB 2;  Length 2946;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2327 TGCCGTGGTGCTTGAGAT 2310


RESULT  10
US-08-461-379A-7/c
; Sequence 7, Application US/08461379A
; Patent No. 5871961
; GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5 Polypeptide,
;    TITLE OF INVENTION:  Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:                (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,379A
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108; 08/104,736
;      APPLICATION NUMBER:  & 07/796,066
;      FILING DATE:  27-OCT-1994; 10-AUG-1993 & 20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-070
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)470-0700
;      TELEFAX:  (610)470-0701
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2946 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  215..2503
US-08-461-379A-7


  Query Match             1.1%;  Score 18;  DB 2;  Length 2946;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2327 TGCCGTGGTGCTTGAGAT 2310


RESULT  11
US-08-462-390B-7/c
; Sequence 7, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR8 Polypeptide, Vector and
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:                (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,390B
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-040
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)407-0700
;      TELEFAX:  (610)407-0701
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2946 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  215..2503
US-08-462-390B-7

  Query Match             1.1%;  Score 18;  DB 2;  Length 2946;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2327 TGCCGTGGTGCTTGAGAT 2310


RESULT  12
US-08-463-074B-7/c
; Sequence 7, Application US/08463074B
; Patent No. 6020155
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR1 Fusion Protein, Vector an
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:                (B) STREET:                444 South Flower St. - Suite 1900
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,074B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38143 (DART-020)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2946 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  215..2503
US-08-463-074B-7

  Query Match             1.1%;  Score 18;  DB 3;  Length 2946;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2327 TGCCGTGGTGCTTGAGAT 2310


RESULT  13
US-08-465-585C-7/c
; Sequence 7, Application US/08465585C
; Patent No. 6027914
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR6 Polypeptide, Vecto
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY,SCHROEDER & POPLAWSKI
;      ADDRESSEE:                (B) STREET:                444South Flower St. - Sui
;      CITY:  Los Angeles
;      STATE:  Californiaa
;      COUNTRY:  USA
;      ZIP:  900071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,585C
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66.38149 (DART-050)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 4894210
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2946 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  215..2503
US-08-465-585C-7

  Query Match             1.1%;  Score 18;  DB 3;  Length 2946;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1333 tgccgtggtgcttgagat 1350
          ||||||||||||||||||
Db   2327 TGCCGTGGTGCTTGAGAT 2310


RESULT  14
US-08-652-446-7/c
```

```
; Sequence 7, Application US/08652446
; Patent No. 6057427
; GENERAL INFORMATION:
;   APPLICANT:  Smith, Kendall A. & Beadling, Carol
;   TITLE OF INVENTION:  Nucleic Acids Encoding CR5
;   TITLE OF INVENTION:  Polypeptide, Vector and Transformed Cell Thereof, and
;   TITLE OF INVENTION:  Expression Thereof
;   NUMBER OF SEQUENCES:  38
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;     ADDRESSEE:          (B) STREET:              444 South Flower St. - Suite 1900
;     CITY:  Los Angeles
;     STATE:  California
;     COUNTRY:  USA
;     ZIP:  90071
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0,
;     SOFTWARE:  Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/652,446
;     FILING DATE:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  EP App. # 96921319.8
;     FILING DATE:  5-JAN-1998
;     APPLICATION NUMBER:  PCT/US/96/09194
;     FILING DATE:  5-JUN-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/330,108
;     FILING DATE:  27-OCT-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/463,074
;     FILING DATE:  5-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/462,337
;     FILING DATE:  5-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/462,390
;     FILING DATE:  5-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/465,585
;     FILING DATE:  5-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/463,081
;     FILING DATE:  5-JUN-1995
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/461,379
;     FILING DATE:  5-JUN-1991
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/739,523
;     FILING DATE:  29-OCT-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Viviana Amzel, Ph. D.
;     REGISTRATION NUMBER:  30,930
;     REFERENCE/DOCKET NUMBER:  FP66 40035
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (213) 622-7700
;     TELEFAX:  (213) 489-4210
; INFORMATION FOR SEQ ID NO:  7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  2946 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  215..2503
US-08-652-446-7

  Query Match             1.1%;  Score 18;  DB 3;  Length 2946;
  Best Local Similarity   100.0%;  Pred. No. 12;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1333 tgccgtggtgcttgagat 1350
           ||||||||||||||||||
Db    2327 TGCCGTGGTGCTTGAGAT 2310


RESULT  15
US-08-308-494A-16/c
; Sequence 16, Application US/08308494A
; Patent No. 5959083
; GENERAL INFORMATION:
;   APPLICANT:  Bosslet, Klaus
;   APPLICANT:  Seeman, Gerhard
;   TITLE OF INVENTION:  Tetravalent Bispecific Receptors, The
;   TITLE OF INVENTION:  Preparation and Use Thereof
;   NUMBER OF SEQUENCES:  24
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;     ADDRESSEE:  Dunner
;     STREET:  1300 I Street, N.W., Suite 700
;     CITY:  Washington
;     STATE:  D.C.
;     COUNTRY:  USA
;     ZIP:  20005-3315
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/308,494A
;     FILING DATE:  21-SEP-1994
;     CLASSIFICATION:  424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 07/891,739
;     FILING DATE:  01-JUN-1992
;     CLASSIFICATION:  424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  DE P4118120.4
;     FILING DATE:  03-JUN-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Kulik, David J.
;     REGISTRATION NUMBER:  36,576
;     REFERENCE/DOCKET NUMBER:  05552-1186-02000
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  202-408-4000
;     TELEFAX:  202-408-4400
; INFORMATION FOR SEQ ID NO:  16:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  300 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
US-08-308-494A-16

  Query Match             1.1%;  Score 17;  DB 2;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      24 tgcagacatgattgctg 40
           |||||||||||||||||
Db      30 TGCAGACATGATTGCTG 14


Search completed: February 15, 2002, 10:46:26
Job time: 1831 sec
```

```
Qy    701 gacttctttaaaaccccctgaagcctacttccaagtgttcgacgatttcgcaggacgcatc 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACTTCTTTAAAACCCCTGAAGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATC 660

Qy    761 accccgaacggcaagctggttgtgtgcctgaacgatcctcacgcagcggagctgggggag 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCCGAACGGCAAGCTGGTTGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAG 720

Qy    821 aggtctgtccgcaagggtatcaagactgttggttacggtaccgctgacgccgtacaggca 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGGTCTGTCCGCAAGGGTATCAAGACTGTTGGTTACGGTACCGCTGACGCCGTACAGGCA 780

Qy    881 caccctgaggttccagcgatggctaccatcgtggattcccaagttgtcgcagaaggcacc 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACCCTGAGGTTCCAGCGATGGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACC 840

Qy    941 cgcgccaccatcaacatcgatggacaggaagtatctgtgattcttcaaatcgctggtgat 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCGCCACCATCAACATCGATGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGAT 900

Qy   1001 cacatggtactcaacggtgcagccgccctgctggccggatacctggtgggtggggacgtc 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CACATGGTACTCAACGGTGCAGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTC 960

Qy   1061 gacaagcttgttgaaggcttgtcggatttctccggcgtgcgacgccgctttgagttccac 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACAAGCTTGTTGAAGGCTTGTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCAC 1020

Qy   1121 ggtgctatcgagggcggcaaatttaatggcgctgctatttatgatgattacgcacaccac 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGTGCTATCGAGGGCGGCAAATTTAATGGCGCTGCTATTTATGATGATTACGCACACCAC 1080

Qy   1181 ccaacggaagtaactgcagtgctcagcgctgcgcgcacccgggtgaaggccgctggaaag 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCAACGGAAGTAACTGCAGTGCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAG 1140

Qy   1241 ggccgtgtcatcgtcgcgttccaaccacatttatactcacgcaccatagaattccaaaag 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCCGTGTCATCGTCGCGTTCCAACCACATTTATACTCACGCACCATAGAATTCCAAAAG 1200

Qy   1301 gagttcgcgggggcactgtcactggcagacgctgccgtggtgcttgagatttacggagcg 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGTTCGCGGGGGCACTGTCACTGGCAGACGCTGCCGTGGTGCTTGAGATTTACGGAGCG 1260

Qy   1361 cgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgatgcgatgaccattcca 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCA 1320

Qy   1421 gtggtgtacgaacctaatttctctgcagtcccagaacgcattgcagaaatcgcaggacct 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCT 1380

Qy   1481 aatgacatcgtgctcaccatgggtgcaggttccgtgaccatgcttgctccagaaatcctg 1540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTG 1440

Qy   1541 gatcagctgcaaaacaat 1558
          ||||||||||||||||||
Db   1441 GATCAGCTGCAAAACAAT 1458
```

```
RESULT   5
E28467
LOCUS       E28467       1458 bp    DNA             PAT       07-FEB-2001
DEFINITION  Novel murC gene.
ACCESSION   E28467
VERSION     E28467.1  GI:13018359
KEYWORDS    JP 1999196876-A/1.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 1458)
  AUTHORS   Takashi,K.H.Y.Y. and Wachi,K.N.
  TITLE     Novel murC gene
  JOURNAL   Patent: JP 1999196876-A 1 27-JUL-1999;
            MITSUBISHI CHEM CORP
COMMENT     OS   Brevibacterium flavum
            PN   JP 1999196876-A/1
            PD   27-JUL-1999
            PF   16-JAN-1998 JP 1998006840
            PR
            PI   TAKASHI KOBAYASHI,HIDEAKI YUKAWA,MASAAKI WACHI,KAZUO NAGAI PC
             C12N15/09,C12N9/00//(C12N15/09,C12R1:13),C12N15/00,(C12N15/00, PC
             C12R1:13)
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   CDS             1. .1458.
FEATURES             Location/Qualifiers
     source          1. .1458
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
BASE COUNT      318 a    420 c    402 g    318 t
ORIGIN
```

```
  Query Match           91.4%;  Score 1445.2;  DB 6;  Length 1458;
  Best Local Similarity  99.5%;  Pred. No. 0;
  Matches 1450;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    101 gtgaccactccacacttggattctgcacaagatattgatctgtcccgcgtccacctcatc 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACCACTCCACACTTGGATTCTGCACAAGATATTGATCTGTCCCGCGTCCACCTCATC 60

Qy    161 ggtattggcggagccggaatgtctggcgttgcccgaatcctgcttgcccgcggtaagaca 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTATTGGCGGAGCCGGAATGTCTGGCGTTGCCCGAATCCTGCTTGCCCGCGGTAAGACA 120

Qy    221 gtcactggttccgatgccaaagattcccgcaccttgcttccactccgcgccgtgggagcc 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTCACTGGTTCCGATGCCAAAGATTCCCGCACCTTGCTTCCACTCCGCGCCGTGGGAGCC 180

Qy    281 accatcgcagtgggacacgctgcggaaaaccttgagctttccggcgaacttcccaccgtc 340
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCATCGCAGTGGGACACGCCGCGGAAAACCTTGAGCTTTCCGGCGAACTTCCCACCGTC 240

Qy    341 gtggtgacctcttttgccgccattccgcaagacaacccggaaccttgttcgtgcacgtgaa 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGGTGACCTCTTTTGCCGCCATTCCGCAAGACAACCCGGAACTTGTTCGTGCACGTGAA 300

Qy    401 gaaggcattccggttattcgtcgctccgatctgttgggcgaattgctggaaggctccacc 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAAGGCATTCCGGTTATTCGTCGCTCCGATCTGTTGGGCGAATTGCTGGAAGGCTCCACC 360

Qy    461 caggtcttgatcgcgggtacccacggtaagacctccaccacctctatgtctgtggtagct 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAGGTCTTGATCGCGGGTACCCACGGTAAGACCTCCACCACCTCTATGTCTGTGGTAGCT 420

Qy    521 atgcaggcagcgggcatggatccaagctttgctatcggcggacagctcaacaaggctggc 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ATGCAGGCAGCGGGCATGGATCCAAGCTTTGCTATCGGCGGACAGCTCAACAAGGCTGGC 480

Qy    581 accaatgcgcaccatggaactggtgaggtctttatcgctgaagcagatgaatctgacgca 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCAATGCGCACCATGGAACTGGTGAGGTCTTTATCGCTGAAGCAGATGAATCTGACGCA 540

Qy    641 tcgctgctgcgctacaagccaaatgttgcagtggtcaccaatgtggaaccagaccacctg 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCGCTGCTGCGCTACAAGCCAAATGTTGCAGTGGTCACCAATGTGGAACCAGACCACCTG 600

Qy    701 gacttctttaaaaccccctgaagcctacttccaagtgttcgacgatttcgcaggacgcatc 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACTTCTTTAAAACCCCTGAAGCCTACTTCCAAGTGTTCGACGATTTCGCAGGACGCATC 660
```

```
Qy    761 accccgaacggcaagctggttgtgtgcctgaacgatcctcacgcagcggagctgggggag 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCCGAACGGCAAGCTGGTTGTGTGCCTGAACGATCCTCACGCAGCGGAGCTGGGGGAG 720

Qy    821 aggtctgtccgcaagggtatcaagactgttggttacggtaccgctgacgccgtacaggca 880
          ||||||||||||||||||||||||||||||||||| |||||||||||||| |||||||||
Db    721 AGGTCTGTCCGCAAGGGTATCAAGACTGTTGGTTATGGTACCGCTGACGCAGTACAGGCA 780

Qy    881 caccctgaggttccagcgatggctaccatcgtggattcccaagttgtcgcagaaggcacc 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CACCCTGAGGTTCCAGCGATGGCTACCATCGTGGATTCCCAAGTTGTCGCAGAAGGCACC 840

Qy    941 cgcgccaccatcaacatcgatggacaggaagtatctgtgattcttcaaatccctggtgat 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCGCCACCATCAACATCGATGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGAT 900

Qy   1001 cacatggtactcaacggtgcagccgccctgctggccggatacctggtgggtggggacgtc 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CACATGGTACTCAACGGTGCAGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTC 960

Qy   1061 gacaagcttgttgaaggcttgtcggatttctccggcgtgcgacgccgctttgagttccac 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GACAAGCTTGTTGAAGGCTTGTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCAC 1020

Qy   1121 ggtgctatcgagggcggcaaatttaatggcgctgctatttatgatgattacgcacaccac 1180
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1021 GGTGCTATCGAGGGCGGCAAATTTAATGGCGCGGCTATTTATGATGATTACGCACACCAC 1080

Qy   1181 ccaacggaagtaactgcagtgctcagcgctgcgcgcacccgggtgaaggccgctggaaag 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCAACGGAAGTAACTGCAGTGCTCAGCGCTGCGCGCACCCGGGTGAAGGCCGCTGGAAAG 1140

Qy   1241 ggccgtgtcatcgtcgcgttccaaccacatttatactcacgcaccatagaattccaaaag 1300
          |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db   1141 GGCCGTGTCATCGTCGCGTTCCAACCACATTTGTACTCACGCACCATGGAATTCCAAAAG 1200

Qy   1301 gagttcgcgggggcactgtcactggcagacgctgccgtggtgcttgagatttacggagcg 1360
          |||||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1201 GAGTTCGCGGAGGCACTGTCACTGGCAGACGCTGCCGTGGTGCTCGAGATTTACGGAGCG 1260

Qy   1361 cgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgatgcgatgaccattcca 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCGATGACCATTCCA 1320

Qy   1421 gtggtgtacgaacctaatttctctgcagtcccagaacgcattgcagaaatcgcaggacct 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGACCT 1380

Qy   1481 aatgacatcgtgctcaccatgggtgcaggttccgtgaccatgcttgctccagaaatcctg 1540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATCCTG 1440

Qy   1541 gatcagctgcaaaacaat 1558
          ||||||||||||||||||
Db   1441 GATCAGCTGCAAAACAAT 1458
```

```
RESULT   6
AE007068/c
LOCUS       AE007068    20400 bp    DNA             BCT       27-APR-2001
DEFINITION  Mycobacterium tuberculosis CDC1551, section 154 of 280 of the
            complete genome.
ACCESSION   AE007068 AE000516
VERSION     AE007068.1  GI:13881884
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis CDC1551.
  ORGANISM  Mycobacterium tuberculosis CDC1551
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 20400)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Whole genome comparison of Mycobacterium tuberculosis clinical and
            laboratory strains
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 20400)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .20400
                     /organism="Mycobacterium tuberculosis CDC1551"
                     /strain="CDC1551"
                     /db_xref="taxon:83331"
                     /note="clinical strain"
     tRNA            221. .309
                     /gene="tRNA-Leu-4"
                     /product="tRNA-Leu"
     gene            221. .309
                     /gene="tRNA-Leu-4"
     gene            complement(427. .744)
                     /gene="MT2200"
     CDS             complement(427. .744)
                     /gene="MT2200"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK46484.1"
                     /db_xref="GI:13881885"
                     /translation="MTRRLRVHNGVEDDLFEAFSYYADAAPDQIDRLYNLFVDAVTRR
                     IPQAPNAFAPLFKHYRHIYLRPFRYYVAYRTTDEAIDILAVRHGMENPNAVEAEISGR
                     TFE"
     gene            complement(741. .956)
                     /gene="MT2201"
     CDS             complement(741. .956)
                     /gene="MT2201"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK46485.1"
                     /db_xref="GI:13881886"
                     /translation="MVVNRALLASVDALSRDEQIELVEHINGNLAEGMHISEANQALI
                     EARANDTDDAHWSTIDDFDKRIRARLG"
     gene            1211. .2269
                     /gene="MT2202"
     CDS             1211. .2269
                     /gene="MT2202"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK46486.1"
                     /db_xref="GI:13881887"
                     /translation="MEAPPYAGDPTFERLRRSFQPADLLPELQAAGVHYTIAVEAADD
                     PAENESLLATARHHDWIARVIGWVPLADPDEVTESSTHGRHRPDASWRRDLRCPGLLP
                     PGCHQPVLVVGLVGQQPEMRPMNPPSGFLRRTPTRRPRDRRDAGRVLADELASYRGRD
                     RLLVLGLARGGVPVGWEVASALGAELDVFLVRKLGVPQWRELAMGALASGGGVVMNDD
                     VVSSLRITDQQVRAAIDSETAELQRRELAYRGGRPVVDPRARIVILVDDGIATGASML
                     AAVRTIRATGPESIVVAVPVGPATACRELAAEADDVVCATMPAAPEAVGQVYNDPHQV
                     TDDEVRELLATPTTGAAT"
     gene            complement(2399. .2755)
                     /gene="MT2203"
     CDS             complement(2399. .2755)
                     /gene="MT2203"
                     /note="identified by Glimmer2; putative"
```

```
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK46487.1"
                     /db_xref="GI:13881888"
                     /translation="MLIIALVLALIGLLALVFAVVTSNQLVAWVCIGASVLGVALLIV
                     DALRERQQGADEADGAGETGVAEEADVDYPEEAPEESQAVDAGVIGSEEPSEEASEA
                     TEESAVSADRSDDSAK"
gene                 complement(2850. .3632)
                     /gene="MT2204"
CDS                  complement(2850. .3632)
                     /gene="MT2204"
                     /note="similar to GB:X77129 SP:P46816 PID:453174
                     PID:2104333 GB:AL123456; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="antigen 84"
                     /protein_id="AAK46488.1"
                     /db_xref="GI:13881889"
                     /translation="MPLTPADVHNVAPSKPPIGKRGYNEDEVDAPLDLVENELTRLIE
                     ENSDLRQRINELDQELAAGGGAGVTPQATQAIPAYEPEPGKPAPAAVSAGMNEEQALK
                     AARVLSLAQDTADRLTNTAKAESDKMLADARANAEQILGEARHTADATVAEARQRADA
                     MLADAQSRSEAQLRQAQEKADALQADAERKHSEIMGTINQQRAVLEGRLEQLRTFERE
                     YRTRLKTYLESQLEELGQRGSAAPVDSNADAGGFDQPNRGKN"
gene                 complement(3900. .4286)
                     /gene="MT2205"
CDS                  complement(3900. .4286)
                     /gene="MT2205"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK46489.1"
                     /db_xref="GI:13881890"
                     /translation="MGTPPESVGWRCRPHRWRRNISSSSSVRSGLSLVVFFQILGFAL
                     FIFWLLLIARVVVEFIRSFSRDWRPTGVTVVILEIIMSITDPPVKVLRRLIPQLTIGA
                     VRFDLSIMVLLLVAFIGMQLAFGAAA"
gene                 complement(4352. .5077)
                     /gene="MT2206"
CDS                  complement(4352. .5077)
                     /gene="MT2206"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK46490.1"
                     /db_xref="GI:13881891"
                     /translation="MNSHCSHTFITDNRSPRARRGHAMSTLHKVKAYFGMAPMEDYDD
                     EYYDDRAPSRGYARPRFDDDYGRYDGRDYDDARSDSRGDLRGEPADYPPPGYRGGYAD
                     EPRPRPREFDRAEMTRPRPGSWLRNSTRGALAMDPRRMAMMPEDGHPLSKITTLRPKD
                     YSEARTIGERPRDGSPVIMDLVSMDNADAKRLVDFAAGLAPALRGSFDKVATKVPLLS
                     PADVDVSPEERRRRIAETGFYAYQ"
gene                 complement(5074. .5868)
                     /gene="MT2207"
CDS                  complement(5074. .5868)
                     /gene="MT2207"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK46491.1"
                     /db_xref="GI:13881892"
                     /translation="MDGMTAMAADLSAYPDRESELTHALAAMRSRLAAAAEAAGRNVG
                     EIELLPITKFFPATDVAILFRLGCRSVGESREQEASAKMAELNRLLAAAELGHSGGVH
                     WHMVGRIQRNKAGSLARWAHTAHSVDSSRLVTALDRAVVAALAEHRRGERLRVYVQVS
                     LDGDGSRGGVDSTTPGAVDRICAQVQESEGLELVGLMGIPPLDWDPDEAFDRLQSEHN
                     RVRAMFPHAIGLSAGMSNDLEVAVKHGSTCVRVGTALLGPRRLRSP"
gene                 complement(5856. .6662)
                     /gene="MT2208"
CDS                  complement(5856. .6662)
                     /gene="MT2208"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAK46492.1"
                     /db_xref="GI:13881893"
                     /translation="MTTTMSTCRPSCAAESIGLLASTRHIARGDTGNVSVRIRRVTTT
                     RAGGVSAPPFDTFNLGDHVGDDPAAVAANRARLAAAIGLPGNRVVWMNQVHGDRVELV
                     DQPRNTALDDTDGLVTATPRLALAVVTADCVPVLMADARAGIAAAVHAGRAGAQRGVV
                     VRALEVMLSLGAQVRDISALLGPAVSGRNYEVPAAMADEVEAALPGSRTTTAAGTPGV
                     DLRAGIACQLRDLGVESIDVDPRCTVADPTLFSHRRDAPTGRPASLVWME"
gene                 complement(6619. .7824)
                     /gene="MT2209"
CDS                  complement(6619. .7824)
                     /gene="MT2209"
                     /note="identified by match to TIGR  protein family HMM
                     TIGR00065"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAK46493.1"
                     /db_xref="GI:13881894"
                     /translation="MRLRVCQQTHRRTEPHQGGSRTMTPPHNYLAVIKVVGIGGGGVN
                     AVNRMIEQGLKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRKAAE
                     DAKDEIEELLRGADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGK
                     RRSNQAENGIAALRESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVQGITD
                     LITTPGLINVDFADVKGIMSGAGTALMGIGSARGEGRSLKAAEIAINSPLLEASMEGA
                     QGVLMSIAGGSDLGLFEINEAASLVQDAAHPDANIIFGTVIDDSLGDEVRVTVIAAGP
                     DVSGPGRKPVMGETGGAHRIESAKAGKLTSTLFEPVDAVSVPLHTNGATLSIGGDDDD
                     VDVPPFMRR"
gene                 complement(7931. .8875)
                     /gene="MT2210"
CDS                  complement(7931. .8875)
                     /gene="MT2210"
                     /note="similar to GP:1353241; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsQ"
                     /protein_id="AAK46494.1"
                     /db_xref="GI:13881895"
```

Query Match 24.8%; Score 392.2; DB 1: Length 20400;
Best Local Similarity 56.4%; Pred. No. 2.2e-84;
Matches 792; Conservative 0; Mismatches 608; Indels 5; Gaps 3;

```
Qy    137 gatctgtcccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccga 196
          ||||||   || || ||| |  |||| || |||||||| || |||| ||| | ||||||
Db  10332 GATCTGCGGCGGGTGCACATGGTCGGCATCGGCGGAGCTGGCATGTCGGGCATCGCCCGA 10273

Qy    197 atcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttg 256
          |||||||| | ||||||    |   ||| | || || || ||||| || || |||    ||
Db  10272 ATCCTGCTGGACCGCGGCGGGCTGGTCTCCGGGTCAGACGCCAAGGAGTCGCGCGGTGTG 10213

Qy    257 cttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgag 316
          | | | || || ||   ||| ||    |||    | |||||||  |||     || ||
Db  10212 CATGCGCTGCGGGCGCGGGGCGCGTTGATCCGGATCGGACACGACGCGTCGTCGCTGGAC 10153

Qy    317 ctttccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaac 376
          || |    ||      |||| |  || || ||  |  ||||||||| ||  ||  ||||
Db  10152 CTGTTGCCCGGTGGCGCCACGGCGGTCGTCACTACCCATGCCGCCATCCCCAAAACCAAC 10093

Qy    377 ccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttg 436
          || || || ||    || |        |||||||| ||  | |  ||  | |   || ||
Db  10092 CCCGAGCTCGTCGAAGCGAGGCGCCGCGGCATTCCCGTGGTGCTGCGGCCGGCCGTGCTG 10033

Qy    437 ggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctcc 496
          | | | ||| |||  ||  ||||    |  || || | || || ||||| |||||  |
Db  10032 GCCAAGTTGATGGCCGGGCGCACCACATTGATGGTCACCGGCACGCACGGCAAGACAACG 9973

Qy    497 accacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatc 556
          || || || |||    | || ||  |||||          ||  | || ||   ||||||  ||
Db  9972 ACGACGTCCATGCTGATCGTCGCCCTGCAGCACTGCGGGCTTGACCCGTCCTTTGCGGTC 9913
```

```
Qy    557 ggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatc 616
Db   9912 GGCGGTGAGCTGGGGAGGCCGGTACCAACGCCCATCACGGCAGTGGCGACTGTTTCGTC 9853
Qy    617 gctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtc 676
Db   9852 GCCGAAGCCGACGAAAGCGATGGCTCGCTGTTGCAGTACACACCCCACGTCGCGGTGATC 9793
Qy    677 accaatgtggaaccagaccacctggacttctttaaaacccctgaagcctacttccaagtg 736
Db   9792 ACCAACATCGAGTCCGATCACCTGGACTTCTACGGCAGCGTCGAGGCGTATGTTGCGGTG 9733
Qy    737 ttcgacgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgat 796
Db   9732 TTCGACTCCTTCGTGGAGCGCATTGTCCCCGGGGGTGCGCTGGTGGTGTGCACTGACGAC 9673
Qy    797 cctcacgcagcggagctgggggagaggtctgtccgcaagggtatcaagactgttggttac 856
Db   9672 CCCGGAGGGGCCGCGCTGGCTCAGCGCGC-GACTGAGCTGGGAATTCGAGTGCTGCGATA 9614
Qy    857 ggtaccgctgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggat 916
Db   9613 CGGGTCGGTGCCGGGTGAGACCATGGCAGCCACGTTGGTCTCGTGGCAGCAACAGGGGGT 9554
Qy    917 tcccaagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatct 976
Db   9553 CGGCGCGGTCGCACATATCCGGTTGGCCTCAGAACTAGCCACAGCACAGGGTCCCCGCGT 9494
Qy    977 g-tgattcttcaaatccctggtgatcacatggtactcaacggtgcagccgccctgctggc 1035
Db   9493 GATGCGGCTGTCGGTGCCCGGGCGACACATGGCGCTCAACGCGCTGGGAGCGCTGCTGGC 9434
Qy   1036 cggatacctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccgg 1095
Db   9433 CGCGGTGCAGATCGGCGCCCCGGCCGACGAGGTGCTCGACGGGCTGGCCGGCTTCGAAGG 9374
Qy   1096 cgtgcgacgccgctttgagttccacggtgctatcgagggcggcaaatttaatggcgctgc 1155
Db   9373 AGTGCGGCGACGATTCGAACTGGTTGGGACCTGCGGCGTCGGAAAGGCGTCGGTGCGCG- 9313
Qy   1156 tatttatgatgattacgcacaccacccaacggaagtaactgcagtgctcagcgctgcgcg 1215
Db   9314 --TGTTCGATGACTACGCCCACCACCCGACGGAGATCAGCGCGACACTGGCGGCGGCGCG 9257
Qy   1216 cacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacatttata 1275
Db   9256 CATGGTGCTCGAACAGGGCGACGGTGGCCGCTGCATGGTTGTGTTTCAACCCCATTTGTA 9197
Qy   1276 ctcacgcaccatagaattccaaaaggagttcgcgggggcactgtcactggcagacgctgc 1335
Db   9196 TTCGCGGACAAAGGCATTCGCTGCTGAGTTTGGGCGTGCGCTGAATGCCGCTGACGAGGT 9137
Qy   1336 cgtggtgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaat 1395
Db   9136 GTTCGTACTCGACGTCTACGGAGCTCGTGAACAACCGCTGGCCGGTGTCAGCGGAGCCAG 9077
Qy   1396 catcaccgatgcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccaga 1455
Db   9076 CGTCGCTGAGCACGTCACTGTGCCGATGCGCTACGTCCCGGATTTTTCGGCGGTCGCACA 9017
Qy   1456 acgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgt 1515
Db   9016 GCAAGTGGCCGCCGCCGCTAGTCCGGGCGACGTCATCGTCACGATGGTGCCGGAGACGT 8957
Qy   1516 gaccatgcttgctccagaaatcctg 1540
Db   8956 GACCTTGCTGGGCCCGGAAATCCTG 8932
```

RESULT 7
MTCY270/c

```
LOCUS       MTCY270      37586 bp    DNA             BCT       03-AUG-2001
DEFINITION  Mycobacterium tuberculosis H37Rv complete genome; segment 96/162.
ACCESSION   Z95388 AL123456
VERSION     Z95388.1  GI:3261759
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis H37Rv.
  ORGANISM  Mycobacterium tuberculosis H37Rv
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 37586)
  AUTHORS   Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
            Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
            Tekaia,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
            Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
            Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
            Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
            Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
            Squares,S., Sqares,R., Sulston,J.E., Taylor,K., Whitehead,S. and
            Barrell,B.G.
  TITLE     Deciphering the biology of Mycobacterium tuberculosis from the
            complete genome sequence
  JOURNAL   Nature. 393 (6685), 537-544 (1998)
  MEDLINE   98295987
REFERENCE   2  (bases 1 to 37586)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
            tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
            Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
            Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
            75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
COMMENT     On Jun 27, 1998 this sequence version replaced gi:2104310.
            Notes:
            Details of M. tuberculosis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
            been renumbered from the original cosmid submissions but the old
            gene designations are in brackets after the new gene numbers.
            Gene prediction was based on a Hidden Markov Model of TB genes
            implemented in TBparse (Krogh) supplemented with visual inspection
            of positional base preference in codons, especially where there is
            an increase in the observed/expected third position G + C.
            CAUTION:  In some cases we may not have predicted the correct
            initiation codon.  Where possible we choose an initiation codon
            (atg, gtg, or ttg) which is preceded by an upstream ribosome
            binding site sequence (optimally 5-13bp before the initiation
            codon).  If this cannot be identified we choose the most upstream
            initiation codon.
FEATURES             Location/Qualifiers
     source          1. .37586
                     /organism="Mycobacterium tuberculosis H37Rv"
                     /strain="H37Rv"
                     /db_xref="taxon:83332"
                     /clone="Y270"
     gene            complement(51. .854)
                     /gene="cysQ"
     CDS             complement(51. .854)
                     /gene="cysQ"
                     /note="Rv2131c, (MTCY270.37), len: 267. cysQ, equivalent
                     to M. leprae CYSQ_MYCLE P46726 cysq protein homolog. (289
                     aa). Contains inositol monophosphatase family signature 1
                     (PS00629), significance uncertain. FASTA best: CYSQ_MYCLE
                     P46726 cysq protein homolog. (289 aa) opt: 1374 z-score:
                     1597.0E(): 0; (77.3% identity in 264 aa overlap)"
                     /codon_start=1
                     /transl_table=11
                     /product="cysQ"
                     /protein_id="CAB08639.1"
                     /db_xref="GI:2104347"
                     /db_xref="SWISS-PROT:O06244"
                     /translation="MVSPAAPDLTDDLTDAELAADLAADAGKLLLQVRAEIGFDQPWT
                     LGEAGDRQANSLLLRRLQAERPGDAVLSEEAHDDLARLKSDRVWIIDPLDGTREFSTP
                     GRDDWAVHIALWRRSSNGQPEITDAAVALPARGNVVYRTDTVTSGAAPAGVPGTLRIA
                     VSATRPPAVLHRIRQTLAIQPVSIGSAGAKAMAVIDGYVDAYLHAGGQWEWDSAAPAG
```

```
                VMLAAGMRASRLDGSPLRYNQLDPYLPDLLMCRAEVAPILLGAIADAWR"
misc_feature    complement(552. .593)
                /gene="cysQ"
                /note="PS00629 Inositol monophosphatase family signature
                1"
gene            945. .1175
                /gene="Rv2132"
CDS             945. .1175
                /gene="Rv2132"
                /note="Rv2132, (MTCY270.36c), len: 76. Function unknown
                but some similarity to Mycobacterium tuberculosis protein
                AL021924|MTV020_4 (84 aa). FASTA score: opt: 142 z-score:
                210.0 E(): 0.00036;    41.8% identity in 55 aa overlap"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein Rv2132"
                /protein_id="CAB08638.1"
                /db_xref="GI:2104346"
                /db_xref="SPTREMBL:O06243"
                /translation="MRTTVSLADDVAAAVQRLRKERSIGLSEAVNELIRAGLTKRQVA
                NRFQQQTYDMGEGIDYSNIGDAIETLDGPASG"
gene            complement(1385. .2173)
                /gene="Rv2133c"
CDS             complement(1385. .2173)
                /gene="Rv2133c"
                /note="Rv2133c, (MTCY270.35), len: 262. Function: unknown
                but equivalent to hypothetical M. leprae protein, Q49774.
                FASTA best: Q49774 B2126_C1_150. (262 aa) opt: 1447
                z-score:1647.5 E(): 0; (79.0% identity in 262 aa overlap)"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein Rv2133c"
                /protein_id="CAB08660.1"
                /db_xref="GI:2104345"
                /db_xref="SPTREMBL:O06242"
                /translation="MLADGELTVLGRIRSASNATFLCESTLGLRSLHCVYKPVSGERP
                LWDFPDGTLAGRELSAYLVSTQLGWNLVPHTIIRDGPAGIGMLQLWVQQPGDAVDSDP
                LPGPDLVDLFPAHRPRPGYLPVLRAYDYAGDEVVLMHADDIRLRRMAVFDVLINNADR
                KGGHILCGIDGQVYGVDHGLCLHVENKLRTVLWGWAGKPIDDQILQAVAGLADALGGP
                LAEALAGRIAAAEIGALRRRAQSLLDQPVMPGPNGHRPIPWPAP"
gene            complement(2184. .2771)
                /gene="Rv2134c"
CDS             complement(2184. .2771)
                /gene="Rv2134c"
                /note="Rv2134c, (MTCY270.34), len: 195. Function: unknown
                but equivalent to hypothetical M. leprae protein, Q49789.
                FASTA best: Q49789 B2126_C3_228, opt: 1192 z-score: 1493.1
                E(): 0 (91.1% identity in 192 aa overlap)"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein Rv2134c"
                /protein_id="CAB08659.1"
                /db_xref="GI:2104344"
                /db_xref="SPTREMBL:O06241"
                /translation="MARAIHVFRTPDRFVAGTVGQPGNRTFYLQAVHDSRVVSVVLEK
                QQVAVLAERIGALLPEVNRRFGTPVPPEPTEIDDLSPLIMPVDAEFRVGTNGLGWDSE
                AQSVVVELLAVTDAEFDASVVLDDTEEGPDAVRVFLTPESARQFATRSYRVISAGRPP
                CPLCDEPLDPEGHICARTNGYRRDVLLGSGDDPAG"
RBS             complement(2776. .2780)
                /note="possible RBS, GGAGG, for Rv2134c"
gene            complement(2835. .3545)
                /gene="Rv2135c"
CDS             complement(2835. .3545)
                /gene="Rv2135c"
                /note="Rv2135c, (MTCY270.33), len: 236. Function: unknown
                but equivalent to hypothetical M. leprae protein, Q49773.
                FASTA best: Q49773 B2126_C1_148 opt: 1183 z-score: 1390.9
                E(): 0; (74.8% identity in 250 aa overlap), also similar
                in C-terminus to PMG2_ECOLI P36942 probable
                phosphoglycerate mutase 2 (215 aa) fasta scores; opt:  212
                z-score: 257.4 E():2.5e-07 27.9% identity in 190 aa
                overlap"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein Rv2135c"
                /protein_id="CAB08658.1"
                /db_xref="GI:2104343"
                /db_xref="SPTREMBL:O06240"
                /translation="MTVILLRHARSTSNTAGVLAGRSGVDLDEKGREQATGLIDRIGD
                LPIRAVASSPMLRCQRTVEPLAEALCLEPLIDDRFSEVDYGEWTGRKIGDLVDEPLWR
                VVQAHPSAAVFPGGEGLAQVQTRAVAAVREHDRRLADQHGHDVLWLACTHGDVIKAVI
                ADAFGMHLDSFQRITADPGSVSVVRYTQLRPFVLHVNHTGARLAPALQAAASAQGASP
                EPNAAVPPGDAVIGGSTD"
gene            complement(3542. .4372)
                /gene="Rv2136c"
CDS             complement(3542. .4372)
                /gene="Rv2136c"
                /note="Rv2136c, (MTCY270.32), len: 276. Function: unknown
                possible membrane protein very similar to hypothetical M.
                leprae protein Q49783. FASTA best: Q49783 B2126_C2_190.
                opt:1023 z-score: 1259.2 E(): 0; (82.4% identity in 187 aa
                overlap) similar to BACA_ECOLI P31054 bacitracin
                resistance protein (273 aa) opt: 477 z-score: 590.1 E():
                7e-26, (35.6% identity in 267 aa overlap)"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein Rv2136c"
                /protein_id="CAB08657.1"
                /db_xref="GI:2104342"
                /db_xref="SPTREMBL:O06239"
                /translation="MSWWQVIVLAAAQGLTEFLPVSSSGHLAIVSRIFFSGDAGASPT
                AVSQLGTEAAVVIYFARDIVRILSANLHGLVVKAHRNTDYRLGWYVIIGTIPICILGL
                FFKDDIRSGVRNLWVVVVTALVVFSGVIALAEYVGRQSRHIERLTWRDAVVVGIAQTLA
                LVPGVSRSGSTISAGLFLGLDRELAARFGFLLAIPAVFASGLFSLPDAFHPVTEGMSA
                TGPQLLVATLIAFVLGLTAVAWLLRFLVRHNMYWFVGYRVLVGTGMLVLLATGTVAAT
                "
gene            complement(4436. .4849)
                /gene="Rv2137c"
CDS             complement(4436. .4849)
                /gene="Rv2137c"
                /note="Rv2137c, (MTCY270.31), len: 137. Function: unknown
                C-terminus is very similar to hypothetical M. leprae
                protein B2126_C2_188 (150 aa). FASTA best: Q49782
                B2126_C2_188. (150 aa) opt: 469 z-score: 623.7 E():
                9.6e-28; (77.2% identity in 101 aa overlap)"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein Rv2137c"
                /protein_id="CAB08656.1"
                /db_xref="GI:2104341"
                /db_xref="SPTREMBL:O06238"
                /translation="MRNMKSTSHESESGKLLSISSCRPREMVLQRYSLGMTVTADRHL
                ADKREEFAVEDISTGIFASGYGQVGDGRSFSFPHIEHRSLVVEIYRPRVAGPVPQAEDV
                VAMAVRGLVDIDLTDERSLAAAVRDSVASAAPVSR"
gene            4864. .5940

  Query Match           24.8%;  Score 392.2;  DB 1;  Length 37586;
  Best Local Similarity  56.4%;  Pred. No. 2.1e-84;
  Matches  792;  Conservative    0;  Mismatches  608;  Indels    5;  Gaps    3;

Qy    137 gatctgtcccgcgtccacctcatcggtattggcggaqccggaatgtctggcgttgcccga 196
          ||||||   || || ||| |  |||| || |||||||| || ||||| ||| | ||||||
Db  19632 GATCTGCGGCGGGTGCACATGGTCGGCATCGGCGGAGCTGGCATGTCGGGCATCGCCCGA 19573

Qy    197 atcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttg 256
          |||||||| | ||||||    |   ||| | || || || ||||| || || |||    ||
Db  19572 ATCCTGCTGGACCGCGGCGGGCTGGTCTCCGGGTCAGACGCCAAGGAGTCGCGCGGTGTG 19513

Qy    257 cttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgag 316
          | | | || || ||   ||| ||    |||    | |||||||  |||        || ||
Db  19512 CATGCGCTGCGGGCGCGGGGCGCGTTGATCCGGATCGGACACGACGCGTCGTCGCTGGAC 19453

Qy    317 ctttccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaac 376
          || |   ||      |||| | || || ||  |   |||||||| || ||  |||| |||
Db  19452 CTGTTGCCCGGTGGCGCCACGGCGGTCGTCACTACCCATGCCGCCATCCCCAAAACCAAC 19393
```

```
Qy    377 ccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttg 436
Db  19392 CCCGAGCTCGTCGAAGCGAGGCGCCGCGGCATTCCCGTGGTGCTGCGGCCGGCCGTGCTG 19333

Qy    437 ggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctcc 496
Db  19332 GCCAAGTTGATGGCCGGGCGCACCACATTGATGGTCACCGGCACGCACGGCAAGACAACG 19273

Qy    497 accacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatc 556
Db  19272 ACGACGTCCATGCTGATCGTCGCCCTGCAGCACTGCGGGCTTGACCCGTCCTTTGCGGTC 19213

Qy    557 ggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatc 616
Db  19212 GGCGGTGAGCTGGGGGAGGCCGGTACCAACGCCCATCACGGCAGTGGCGACTGTTTCGTC 19153

Qy    617 gctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtc 676
Db  19152 GCCGAAGCCGACGAAAGCGATGGCTCGCTGTTGCAGTACACACCCCACGTCGCGGTGATC 19093

Qy    677 accaatgtggaaccagaccacctggacttctttaaaaccccctgaagcctacttccaagtg 736
Db  19092 ACCAACATCGAGTCCGATCACCTGGACTTCTACGGCAGCGTCGAGGCGTATGTTGCGGTG 19033

Qy    737 ttcgacgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgat 796
Db  19032 TTCGACTCCTTCGTGGAGCGCATTGTCCCCGGGGGTGCGCTGGTGGTGTGCACTGACGAC 18973

Qy    797 cctcacgcagcggagctgggggagaggtctgtccgcaagggtatcaagactgttggttac 856
Db  18972 CCCGGAGGGGCCGCGCTGGCTCAGCGCGC-GACTGAGCTGGGAATTCGAGTGCTGCGATA 18914

Qy    857 ggtaccgctgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggat 916
Db  18913 CGGGTCGGTGCCGGGTGAGACCATGGCAGCCACGTTGGTCTCGTGGCAGCAACAGGGGGT 18854

Qy    917 tcccaagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatct 976
Db  18853 CGGCGCGGTCGCACATATCCGGTTGGCCTCAGAACTAGCCACAGCACAGGGTCCCCGCGT 18794

Qy    977 g-tgattcttcaaatccctggtgatcacatggtactcaacggtgcagccgccctgctggc 1035
Db  18793 GATGCGGCTGTCGGTGCCCGGGCGACACATGGCGCTCAACGCGCTGGGAGCGCTGCTGGC 18734

Qy   1036 cggatacctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccgg 1095
Db  18733 CGCGGTGCAGATCGGCGCCCCGGCCGACGAGGTGCTCGACGGGCTGGCCGGCTTCGAAGG 18674

Qy   1096 cgtgcgacgccgctttgagttccacggtgctatcgagggcggcaaatttaatggcgctgc 1155
Db  18673 AGTGCGGCGACGATTCGAACTGGTTGGGACCTGCGGCGTCGGAAAGGCGTCGGTGCGCG- 18613

Qy   1156 tatttatgatgattacgcacaccacccaacggaagtaactgcagtgctcagcgctgcgcg 1215
Db  18614 --TGTTCGATGACTACGCCCACCACCCGACGGAGATCAGCGCGACACTGGCGGCGGCGCG 18557

Qy   1216 cacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacatttata 1275
Db  18556 CATGGTGCTCGAACAGGGCGACGGTGGCCGCTGCATGGTTGTGTTTCAACCCCATTTGTA 18497

Qy   1276 ctcacgcaccatagaattccaaaaggagttcgcgggggcactgtcactggcagacgctgc 1335
Db  18496 TTCGCGGACAAAGGCATTCGCTGCTGAGTTTGGGCGTGCGCTGAATGCCGCTGACGAGGT 18437

Qy   1336 cgtggtgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaat 1395
Db  18436 GTTCGTACTCGACGTCTACGGAGCTCGTGAACAACCGCTGGCCGGTGTCAGCGGAGCCAG 18377

Qy   1396 catcaccgatgcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccaga 1455
Db  18376 CGTCGCTGAGCACGTCACTGTGCCGATGCGCTACGTCCCGGATTTTTCGGCGGTCGCACA 18317

Qy   1456 acgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgt 1515
```

```
Db  18316 GCAAGTGGCCGCCGCCGCTAGTCCGGGCGACGTCATCGTCACGATGGGTGCCGGAGACGT 18257

Qy   1516 gaccatgcttgctccagaaatcctg 1540
Db  18256 GACCTTGCTGGGCCCGGAAATCCTG 18232


RESULT   8
AX191745/c
LOCUS       AX191745    37586 bp    DNA             PAT       15-AUG-2001
DEFINITION  Sequence 27 from Patent WO0149775.
ACCESSION   AX191745
VERSION     AX191745.1  GI:15209914
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis.
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 37586)
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 27 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .37586
                     /organism="Mycobacterium tuberculosis"
                     /db_xref="taxon:1773"
BASE COUNT     6232 a   13189 c   12267 g    5898 t
ORIGIN


  Query Match           24.8%;  Score 392.2;  DB 6;  Length 37586;
  Best Local Similarity   56.4%;  Pred. No. 2.1e-84;
  Matches  792;  Conservative    0;  Mismatches  608;  Indels    5;  Gaps    3;

Qy    137 gatctgtcccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccga 196
Db  19632 GATCTGCGGCGGGTGCACATGGTCGGCATCGGCGGAGCTGGCATGTCGGGCATCGCCCGA 19573

Qy    197 atcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttg 256
Db  19572 ATCCTGCTGGACCGCGGCGGGCTGGTCTCCGGGTCAGACGCCAAGGAGTCGCGCGGTGTG 19513

Qy    257 cttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgag 316
Db  19512 CATGCGCTGCGGGCGCGGGGCGCGTTGATCCGGATCGGACACGACGCCGTCGTCGCTGGAC 19453

Qy    317 ctttccggcgaaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaac 376
Db  19452 CTGTTGCCCGGTGGCGCCACGGCGGTCGTCACTACCCATGCCGCCATCCCCAAAACCAAC 19393

Qy    377 ccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttg 436
Db  19392 CCCGAGCTCGTCGAAGCGAGGCGCCGCGGCATTCCCGTGGTGCTGCGGCCGGCCGTGCTG 19333

Qy    437 ggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctcc 496
Db  19332 GCCAAGTTGATGGCCGGGCGCACCACATTGATGGTCACCGGCACGCACGGCAAGACAACG 19273

Qy    497 accacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatc 556
Db  19272 ACGACGTCCATGCTGATCGTCGCCCTGCAGCACTGCGGGCTTGACCCGTCCTTTGCGGTC 19213

Qy    557 ggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatc 616
Db  19212 GGCGGTGAGCTGGGGGAGGCCGGTACCAACGCCCATCACGGCAGTGGCGACTGTTTCGTC 19153

Qy    617 gctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtc 676
Db  19152 GCCGAAGCCGACGAAAGCGATGGCTCGCTGTTGCAGTACACACCCCACGTCGCGGTGATC 19093
```

```
Qy    677 accaatgtggaaccagaccacctggacttctttaaaacccctgaagcctacttccaagtg 736
Db  19092 ACCAACATCGAGTCCGATCACCTGGACTTCTACGGCAGCGTCGAGGCGTATGTTGCGGTG 19033

Qy    737 ttcgacgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgat 796
Db  19032 TTCGACTCCTTCGTGGAGCGCATTGTCCCCGGGGGTGCGCTGGTGGTGTGCACTGACGAC 18973

Qy    797 cctcacgcagcggagctgggggagaggtctgtccgcaagggtatcaagactgttggttac 856
Db  18972 CCCGGAGGGGCCGCGCTGGCTCAGCGCGC-GACTGAGCTGGGAATTCGAGTGCTGCGATA 18914

Qy    857 ggtaccgctgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggat 916
Db  18913 CGGGTCGGTGCCGGGTGAGACCATGGCAGCCACGTTGGTCTCGTGGCAGCAACAGGGGGT 18854

Qy    917 tcccaagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatct 976
Db  18853 CGGCGCGGTCGCACATATCCGGTTGGCCTCAGAACTAGCCACAGCACAGGGTCCCCGCGT 18794

Qy    977 g-tgattcttcaaatccctggtgatcacatggtactcaacggtgcagccgccctgctggc 1035
Db  18793 GATGCGGCTGTCGGTGCCCGGCGACACATGGCGCTCAACGCGCTGGGAGCGCTGCTGGC 18734

Qy   1036 cggatacctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccgg 1095
Db  18733 CGCGGTGCAGATCGGCGCCCCGGCCGACGAGGTGCTCGACGGGCTGGCCGGCTTCGAAGG 18674

Qy   1096 cgtgcgacgccgctttgagttccacggtgctatcgagggcggcaaatttaatggcgctgc 1155
Db  18673 AGTGCGGCGACGATTCGAACTGGTTGGGACCTGCGGCGTCGGAAAGGCGTCGGTGCGCG- 18613

Qy   1156 tatttatgatgattacgcacaccacccaacggaagtaactgcagtgctcagcgctgcgcg 1215
Db  18614 --TGTTCGATGACTACGCCCACCACCCGACGGAGATCAGCGCGACACTGGCGGCGGCGCG 18557

Qy   1216 cacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacatttata 1275
Db  18556 CATGGTGCTCGAACAGGGCGACGGTGGCCGCTGCATGGTTGTGTTTCAACCCCATTTGTA 18497

Qy   1276 ctcacgcaccatagaattccaaaaggagttcgcgggggcactgtcactggcagacgctgc 1335
Db  18496 TTCGCGGACAAAGGCATTCGCTGCTGAGTTTGGGCGTGCGCTGAATGCCGCTGACGAGGT 18437

Qy   1336 cgtggtgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaat 1395
Db  18436 GTTCGTACTCGACGTCTACGGAGCTCGTGAACAACCGCTGGCCGGTGTCAGCGGAGCCAG 18377

Qy   1396 catcaccgatgcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccaga 1455
Db  18376 CGTCGCTGAGCACGTCACTGTGCCGATGCGCTACGTCCCGGATTTTTCGGCGGTCGCACA 18317

Qy   1456 acgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgt 1515
Db  18316 GCAAGTGGCCGCCGCCGCTAGTCCGGGCGACGTCATCGTCACGATGGGTGCCGGAGACGT 18257

Qy   1516 gaccatgcttgctccagaaatcctg 1540
Db  18256 GACCTTGCTGGGCCCGGAAATCCTG 18232
```

```
RESULT    9
MLEPRTN4
LOCUS       MLEPRTN4   348450 bp    DNA             BCT       20-FEB-2001
DEFINITION  Mycobacterium leprae strain TN complete genome; segment 4/10.
ACCESSION   AL583920 AL450380
VERSION     AL583920.1  GI:13092922
KEYWORDS    .
SOURCE      Mycobacterium leprae.
  ORGANISM  Mycobacterium leprae
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (bases 1 to 348450)
  AUTHORS   Cole,S.T., Eiglmeier,K., Parkhill,J., James,K.D., Thomson,N.R.,
            Wheeler,P.R., Honore,N., Garnier,T., Churcher,C., Harris,D.,
            Mungall,K., Basham,D., Brown,D., Chillingworth,T., Connor,R.,
            Davies,R.M., Devlin,K., Duthoy,S., Feltwell,T., Fraser,A.,
            Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Lacroix,C.,
            Maclean,J., Moule,S., Murphy,L., Oliver, Quail,M.A.,
            Rajandream,M.-A., Rutherford,K.M., Rutter,S., Seeger,K., Simon,S.,
            Simmonds,M., Skelton,J., Squares,R., Squares,S., Stevens,K.,
            Taylor,K., Whitehead,S., Woodward,J.R. and Barrell,B.G.
  TITLE     Massive gene decay in the leprosy bacillus
  JOURNAL   Nature 409 (6823), 1007-1011 (2001)
  MEDLINE   21128732
REFERENCE   2  (bases 1 to 348450)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2001) Submitted on behalf of the Mycobacterium
            leprae sequencing teams, The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridge, CB10 1SA, UK Unitie de Genetique
            Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Rouux,
            75724, Paris Cedex, France. E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of M. leprae sequencing at the Sanger Centre are available
            from http://www.sanger.ac.uk/Projects/M_leprae/ A relational
            datbase containing the M. leprae sequences is available from
            http://genolist.pasteur.fr/Leproma/.
FEATURES             Location/Qualifiers
     source          1. .348450
                     /organism="Mycobacterium leprae"
                     /strain="TN"
                     /db_xref="taxon:1769"
     gene            551. .1786
                     /gene="ML0842"
     CDS             551. .1786
                     /gene="ML0842"
                     /note="Similar to Mycobacterium tuberculosis hypothetical
                     44.6 kDa protein Rv1464 or MTV007.11 TR:O53155
                     (EMBL:AL021184) (417 aa) fasta scores: E(): 0, 43.6% id in
                     408 aa and to many other bacterial nifS-homologues, e.g.
                     Bacillus subtilis YurW protein yurW TR:O32164
                     (EMBL:Z99120) (406 aa) fasta scores: E(): 0, 46.7% id in
                     405 aa. Previously sequenced as TR:O32975 (EMBL:Z98741).
                     Contains Pfam match to entry PF00266 aminotran_5,
                     Aminotransferases class-V.
                     Similar to ML0117, ML0596 and ML1708"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAC31223.1"
                     /db_xref="GI:13092923"
                     /translation="MVDVFDIQAVRADPPILQETVNGKPLIWFDNAATTQKPQVVIDR
                     LSYFYAHENSNIHRAAHELAARATDAYEEARETARRFIGAAKAQEIIFVRGTTEAINL
                     VAYAWGGKHLQPGDEVVITHLEHHANIVPWQLLSSQTGAILKVAPVDDAGNLLMSEFE
                     DLLGPRTKLVAATQVSNALGTVTQGEKIVELGHRYGARVLIDGAQSIPHLPINVSELG
                     ADFFVFSGHKIYGPTGIGVLYGCEDVLTEMPPWQGGGNMIVDVTLERSLYQGPPNKFE
                     AGTGNIADAVGLGEALRYVERVGVQRIASHEQALLDYATPRLADIPGVRLVGTATEKA
                     SVLSFVLAGHEPLEVGKALNAEGIAVRAGHHCAQPVLRRLGLEATVRPSFAFYNTYEE
                     IDVFINVVRRIAEGGTNIG"
     misc_feature    653. .1738
                     /gene="ML0842"
                     /note="Pfam match to entry PF00266 aminotran_5,
                     Aminotransferases class-V, score 23.70, E-value 1.7e-08"
     gene            1908. .2395
                     /gene="lppP"
                     /note="ML0843"
     CDS             1908. .2395
                     /gene="lppP"
                     /note="Possible pseudogene of M. tuberculosis orthologue
                     lppP (Best blastx score 294)"
                     /codon_start=1
                     /pseudo
                     /transl_table=11
                     /product="lipoprotein (pseudogene)"
```

```
Qy   841 caagactgttggttacggtaccgctgacgccgtacaggcacaccctgaggttccagcgat 900
         |||||||||||||||||||||  ||||| |||||||||||||||||||||||||||||||||
Db  3887 CAAGACTGTTGGTTACGGTACCCGTGACGCAGTACAGGCACACCCTGAGGTTCCAGCGAT 3946

Qy   901 ggctaccatcgtggattcccaagttgtcgcagaaggcacccgcgccaccatcaacatcga 960
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3947 GGCTACCGTCGTGGATTCCCAAGTTGTCGCAGAAGGCACCCGCGCCACCATCAACATCGA 4006

Qy   961 tggacaggaagtatctgtgattcttcaaatccctggtgatcacatggtactcaacggtgc 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4007 TGGACAGGAAGTATCTGTGATTCTTCAAATCCCTGGTGATCACATGGTACTCAACGGTGC 4066

Qy  1021 agccgccctgctggccggatacctggtgggtggggacgtcgacaagcttgttgaaggctt 1080
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  4067 AGCCGCCCTGCTGGCCGGATACCTGGTGGGTGGGGACGTCGATAAGCTTGTTGAAGGCTT 4126

Qy  1081 gtcggatttctccggcgtgcgacgccgctttgagttccacggtgctatcgagggcggcaa 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4127 GTCGGATTTCTCCGGCGTGCGACGCCGCTTTGAGTTCCACGGTGCTATCGAGGGCGGCAA 4186

Qy  1141 atttaatggcgctgctatttatgatgattacgcacaccacccaacggaagtaactgcagt 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4187 ATTTAATGGCGCTGCTATTTATGATGATTACGCACACCACCCAACGGAAGTAACTGCAGT 4246

Qy  1201 gctcag 1206
         ||||||
Db  4247 GCTCAG 4252
```

```
RESULT   6
AB003132
LOCUS       AB003132     4116 bp    DNA        BCT       04-AUG-1997
DEFINITION  Corynebacterium glutamicum gene for MurC, FtsQ, FtsZ, complete cds.
ACCESSION   AB003132
VERSION     AB003132.1  GI:2308989
KEYWORDS    FtsZ; FtsQ; MurC.
SOURCE      Corynebacterium glutamicum DNA.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 4116)
  AUTHORS   Kobayashi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1997) to the DDBJ/EMBL/GenBank databases. Miki
            Kobayashi, Mitsubishi Chemical Corp., Tsukuba Research Center; Ami
            8-3-1, Inashiki, Ibaraki, Inashiki, Ibaraki 300-03, Japan
            (E-mail:3709292@cc.m-kagaku.co.jp, Tel:0298-87-1011,
            Fax:0298-87-3259)
REFERENCE   2  (sites)
  AUTHORS   Kobayashi,M., Asai,Y., Hatakeyama,K., Kijima,N., Wachi,M., Nagai,K.
            and Yukawa,H.
  TITLE     Cloning, sequencing, and characterization of the ftsZ gene from
            coryneform bacteria
  JOURNAL   [redacted]
  MEDLINE   [redacted]
FEATURES             Location/Qualifiers
     source          1..4116
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            1..273
                     /gene="murC"
     CDS             <1..273
                     /gene="murC"
                     /codon_start=1
                     /transl_table=11
                     /product="MurC"
                     /protein_id="BAA21685.1"
                     /db_xref="GI:2308990"
                     /translation="EFQKEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIP
                     VVYEPNFSAVPERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            301..969
                     /gene="ftsQ"
     CDS             301..969
                     /gene="ftsQ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsQ"
                     /protein_id="BAA21686.1"
                     /db_xref="GI:2308991"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKAQDAQM
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKRDGQRWNISSPSM
                     VTVR"
     gene            1220..2536
                     /gene="ftsZ"
     CDS             1220..2536
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsZ"
                     /protein_id="BAA21687.1"
                     /db_xref="GI:2308992"
                     /translation="MTSPNNYLAKIKVVGVGGGGDNAVNRMIEEGLKGVEFIAVNTDS
                     QALMFSDADVKLDIGREATRGLGAGANPEVGRASAEYHKNEIEETIKGADMVFVTAGE
                     GGGAGTGAAPVVGRIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAALKEVCDTLIV
                     IPNDRLLELGDANLSIMEAFRAADEVLHNGVQGITNLITIPCINVDFADVRSVMSEAG
                     SALMGVGSARGDNRVVSATEQAINSPLLEATMDGATGVLLSFAGGSDLGLMEVNAAAS
                     MVRERSDEDVNLIFGTIIDDNLGDEVRVTVIATGFDAARASAAENRRAGIPAAPAAEP
                     VQQQQVPTTNATLPPEKESIFGGAREENDPYLSRSAGARHRIEETRSGGGLFTTGNDR
                     DYRRDERREDHRDERRDDRSYDRRDDRRDDRRDDRGDDLDVPSFLQ"
     CDS             2564..3304
                     /note="hypothetical protein similar to YPIH"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA21688.1"
                     /db_xref="GI:2308993"
                     /translation="MDSLDPRNRPVRKVFTTRAGGVSQSPYATFNLGDHVGYYPKAVA
                     SNRNRLADIIGLSPDKVVYMEQIHSNTVTVIDEAPADGQAVEATDALVTTQRGLALAV
                     LVADCVPVLLSDTDAGVIAAVHAGRMGARNGIVAKTIAKMEELGAKPSRIHALMGAAA
                     SGAKYEVPEAMARDVEAKLPGSLARTTKGTTGLDIRAGLLRQMLSLGVQMIDSDPRCT
                     IEDEDLFSYRREGTTGRQAGVVWLPKEA"
     CDS             3345..4010
                     /note="hypothetical protein similar to YPI1 or YPT5"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA21689.1"
                     /db_xref="GI:2308994"
                     /translation="MQARIDATLNEHNRPEGSVRLLPVTKFHPVEDIKILQELGVTAV
                     GENREQEARAKALELPDMDFHMIGQIQSKKANSIARWSAAVHSVDSEKIAEALGRGVA
                     LALERGDRTSDELPCFIQLSLDGDPSRGGTPLSQVTQLADCISDTTHLRFEGLMCVPP
                     LGWDPEKAFSQARDVLSGLEEHFDRSLEFSAGMSGDLVAAIKHGSTIVRVGTEILGNR
                     PLA"
BASE COUNT     1009 a   1106 c   1148 g    853 t
ORIGIN

  Query Match            14.9%;  Score 236;  DB 1;  Length 4116;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-119;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy  1346 gagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgat 1405
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    58 GAGATTTACGGAGCGCGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGAT 117

Qy  1406 gcgatgaccattccagtggtgtacgaacctaatttctctgcagtcccagaacgcattgca 1465
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   118 GCGATGACCATTCCAGTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCA 177

Qy  1466 gaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgtgaccatgctt 1525
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   178 GAAATCGCAGGACCTAATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTT 237

Qy  1526 gctccagaaatcctggatcagctgcaaaacaattaggacgtaagtgaacaaggcag 1581
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

QP 501.B43

1.2b

```
RBS             2549. .2555
                /note="possible RBS"
gene            2559. .4112
                /gene="narK"
                /note="ML0844"
CDS             2559. .4112
                /gene="narK"
                /note="Similar to Mycobacterium tuberculosis putative
                nitrite extrusion protein narK1 or Rv2329c or MTCY3G12.05
                TR:P71883 (EMBL:Z79702) (515 aa) fasta scores: E(): 0,
                69.3% id in 488 aa and to Escherichia coli nitrite
                extrusion protein 2 narU SW:NARU_ECOLI (P37758; P77696)
                (462 aa) fasta scores: E(): 0, 36.6% id in 459 aa. There
                is a frameshift near the C-terminus relative to the M.
                tuberculosis homologue. Previously sequenced as TR:O32974
                (EMBL:Z98741). Contains hydrophobic, probable
                membrane-spanning regions."
                /codon_start=1
                /transl_table=11
                /product="putative nitrite extrusion protein"
                /protein_id="CAC31225.1"
                /db_xref="GI:13092924"
                /translation="MDQLVLLQAEESLYGPGRRSKTVPIDHRHSLKSSFFRLRLRRAC
                RISHWDPEDQAAWEAGNKTIARLNLLWSVLTVHLGISVWTLNPVMELFMPKDVTGFSA
                GDKFLLATTATLVGACLRVPTSLATALFGGRKWAIFSVMVLLIPTIATMVLLAHPGLP
                LWPYLACAALTGLGGGNFAASMTNANAFYPHRLKGAAFGLAGGAGNLGVSVIQVVGML
                VIASVGDRKPYIVCGLYVVLLIIAGIGAMLFMNDIEHHRIGVNTIRPVLFVVVSTRDS
                WVLALLYLASPGSPIGFSPAPGQVLETNPVAGGQSTAQAALHAAELAFIGPTLAAVAR
                FWGGRLADRLGGSRVTLVVFGAMVPAAGLLGVLGIIEGSRVCPIRGVMMASYFAGFIT
                LFILSGLGNGSVYKMIPTIFEACSHSLGINDDECRDWSRVISGVVIGFVAEVGALGGV
                GIDLALRESYLNTGGVTAAFWIFMLCYAAAGVLTWKMYVCRPLPGNLHDEAANAFAAS
                VGASRTHRG"
gene            complement(4100. .4681)
                /gene="ML0845"
CDS             complement(4100. .4681)
                /gene="ML0845"
                /note="Unknown function. Similar to part of some acyl-CoA
                oxidases e.g. Arabidopsis thaliana  acyl-CoA oxidase cx2
                TR:O65201 (EMBL:AF057043) (692 aa) fasta scores: E():
                1.5e-06, 35.2% id in 125 aa. Previously sequenced as
                TR:O32973 (EMBL:Z98741) (193 aa) fasta scores: E(): 0,
                99.5% id in 193 aa."
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="CAC31226.1"
                /db_xref="GI:13092925"
                /translation="MPLRPTNAASERVTKRTATETILQRMVGARQDSEKKGSLFNRGT
                QGTIFEKCEDCLLSAVARRRQVKSKEMFAFDTFNTVQDHLLHAANEHTDRVVLETFLA
                GIESCNDPAARKLLDIVCDLYALSVKARHGTSSIDTSPPSALKAIAKGSRPMPSAAAL
                CRDVGRRPSESQSSCVTPRCRTSRICPSELQPR"
gene            complement(4959. .5480)
                /gene="ML0846"
CDS             complement(4959. .5480)
                /gene="ML0846"
                /note="Possible pseudogene similar to M. tuberculosis
                paralogue Rv1747 (Best blastx score 127)"
                /codon_start=1
                /pseudo
                /transl_table=11
                /product="ABC transporter (pseudogene)"
gene            complement(5522. .5848)
                /gene="ML0847"
CDS             complement(5522. .5848)
                /gene="ML0847"
                /note="Possible pseudogene of M. tuberculosis orthologue
                Rv2327 (Best blastx score 184)"
                /codon_start=1
                /pseudo
                /transl_table=11
                /product="conserved hypothetical protein (pseudogene)"
gene            5906. .8080
                /gene="ML0848"
CDS             5906. .8080
                /gene="ML0848"
                /note="Similar to Mycobacterium tuberculosis hypothetical
                ABC transporter ATP-binding protein Rv2326c or MTCY3G12.08
                SW:YN26_MYCTU (P71886) (697 aa) fasta scores: E(): 0,
                76.9% id in 697 aa. Shares similar domains with many
                ABC-type transporters e.g. Streptomyces roseofulvus ATPase
                component of putative ABC transporter frnD TR:O68910
                (EMBL:AF058302) (524 aa) fasta scores: E(): 8.4e-14, 31.5%
                id in 495 aa and Synechococcus sp. nitrate transport
                ATP-binding protein NrtD nrtD SW:NRTD_SYNP7 (P38046) (274
                aa) fasta scores: E(): 1.6e-12, 34.8% id in 198 aa.
                Previously sequenced as TR:O32971 (EMBL:Z98741). Contains
                hydrophobic, possible membrane-spanning regions. Contains
                2 Pfam matches to entry PF00005 ABC_tran, ABC transporter.
                Contains 2 x PS00017 ATP/GTP-binding site motif A
                (P-loop). Contains 2 x PS00211 ABC transporters family
                signature."
                /codon_start=1
                /transl_table=11
                /product="ABC transporter"
                /protein_id="CAC31229.1"
                /db_xref="GI:13092926"
                /translation="MDLCIHRRMALLCQATRRCHRWKTQLGRCAAPDLAPGTGTIGKV
                TAHDPRHLHHRSGSLQPGELAQASVLAALCAVTAIVSVVVPFAAGLALLGTVPMGLLA
                IRYRFRVLMTAMVAAGVIAFLITGLGGPIAVVNSAYIGGLTGVVKRKGQGTLTVIALA
                LFAGLAFGAANVVALVVLGRLRHLIFKAMTANVDGIAATLTWMHLPWVAVQLKRYFAD
                GLQHWPWMLLGYFVITILVVSLIGWHVLSRVLERIRDIPDVHKLDAPSACNEDAPVGP
                VPVWLDKVRFRYPHAGQDALREVSLDLRVGEHVAVTGANGSGKTTLMLILAGREPTSG
                TVDRPGAVGLGKLGGTAVVLQHPESQVLGTRVADDVVWGLPPGTDVDVNRLLREVGLD
                AFAERDTGSLSGGELQRLALAAALAREPSLLIADEVTSMVDRQGRDALLGVLSGLTKR
                HPIALVHITHYNNEADTADRTINLSDSPDNAGMAETVAPPVSTVAVDHRPHAPVLELV
                GVGHEYGSGTPWAKAALHDISFVVRQGDGVLVYGSNGSGKSTLAWIMAGLMVPTTGAC
                LIDGRPTHEHVGAVALSFQAARLQLMRSRVDLEVASAAGFSPRDEDRVAAALGVVGLD
                PALAKRRIDQLSGGQMRRVVLAGLLACSPRALILDEPLAGLDAVSQRGLLRLLEDLRC
                ERGLTVVVTSHDFVGLEDVCPRTVHLRNGALESVSTTAGGTS"
misc_feature    6821. .7336
                /gene="ML0848"
                /note="Pfam match to entry PF00005 ABC_tran, ABC
                transporter, score 112.30, E-value 9e-30"
misc_feature    6842. .6865
                /gene="ML0848"
                /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
misc_feature    7109. .7153
```

```
Query Match           20.8%;  Score 329.2;  DB 1;  Length 348450;
Best Local Similarity  53.5%;  Pred. No. 2.8e-69;
Matches  760;  Conservative    0;  Mismatches  648;  Indels   12;  Gaps    3;

Qy    137 gatctgtcccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccga 196
          ||  ||   || || ||| |  | || ||||| ||| |||| ||||| ||  |||| ||
Db  84545 GAGTTGCAGCGGGTGCACATGGTGGGCATTGGGGGATCCGGGATGTCGGGTATTGCGCGC 84604

Qy    197 atcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttg 256
          |||||| | | ||| ||          ||   |||| || |||||||| || || ||     ||
Db  84605 ATCCTGTTAGACCGGGGCGGTCTGGTGTCTGGCTCAGATGCCAAGGAGTCGCGTGTTGTG 84664

Qy    257 cttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgag 316
          |   |||| || ||    || ||    |||    | ||||| | |||| |   || ||
Db  84665 CACGCACTGCGGGCGCGCGGCGCGCTGATCCGCATCGGACATGATGCGTCATTGCTGGAC 84724

Qy    317 ctttccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaac 376
          |       |      | ||||  ||  | |||       | ||||| ||  |    |||
Db  84725 TTGCTGCCTGGCGGCGCTACCGCAGTCATTACCATCCGCACTGCCATCCCCAAGACTAAC 84784

Qy    377 ccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttg 436
          || || || ||    || ||        ||||||||||| |    || || |   |||||
Db  84785 CCCGAGCTGGTCGAAGCCCGGAGGCGCGGCATTCCGGTTCTATTGCGATCGGCCGTGTTG 84844

Qy    437 ggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctcc 496
          | |    || |||| || |  |||  | |  || | || || || ||||| |||||  |
Db  84845 GCCAGGCTGATGGATGGGTGTACCACGCTGATGGTTGCTGGCACACACGGCAAGACTACG 84904
```

```
Qy     497 accacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatc 556
Db   84905 ACGACGTCGATGCTCGTTGTGGCCCTGCAGCACTGTGGATGTGACCCATCCTTCGTGGTT 84964
Qy     557 ggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatc 616
Db   84965 GGCGGTGAGCTTGCCGTGGTGGGCACCAACGCCCACCACGGCCTCGGCGCCTGCTTCGTT 85024
Qy     617 gctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtc 676
Db   85025 GCCGAGGCCGACGAAAGCGATGGCTCGTTACTGGAATACACGCCCAACGTAGCGGTGGTC 85084
Qy     677 accaatgtggaaccagaccacctggacttctttaaaaccccctgaagcctacttccaagtg 736
Db   85085 ACCAACATCGACTCCGATCACTTGGATTTCTACGGTAGTGTCGATGCTTATATCAGGGTG 85144
Qy     737 ttcgacgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgat 796
Db   85145 TTCGACTCCTTCGTGGAGCGATTCGCACTGGGGGGAGCGCTCGTGGTGTGCAACGACGAT 85204
Qy     797 cctcacgcagcggagctgggggagaggtctgtccgcaagggtatcaagactgttggttac 856
Db   85205 CCTGGGGCGGCGCGCTGGCCCGGCGCACCGCCGAGCTTGGGATCCGGGTGTTGCGCTAC 85264
Qy     857 ggtaccgctgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggat 916
Db   85265 GGCTCCGACGACCGGATAGGTGAGACCTTGGCCGCCAGGTTGTTGTCGTG------GGAG 85318
Qy     917 tcccaagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatct 976
Db   85319 CAACAGGGCACCGGAGCGGTCGCGCACATCCAGTTGGCTGGTCAACCAAACTCACGGGTG 85378
Qy     977 gtgattcttcaaatccctggtgatcacatggtactcaacggtgcagccgccctgctggcc 1036
Db   85379 ATGCGGCTACCGGTACCTGGGCGACACATGGCGCTCAATGCGCTAGGTGCGTTGCTGGCG 85438
Qy    1037 ggatacctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccggc 1096
Db   85439 GCGATCGAGATTGGTGCGTCGACCGATGAAGTGCTCGACGGGCTGGCTGGTTTCAGGGGT 85498
Qy    1097 gtgcgacgccgctttgagttccacggtgctatcgagggcggcaa--atttaatggcgctg 1154
Db   85499 GTGCGTCGCCGCTTCGAACTGGTTGGGACTTCAGGAGTCGGCCAGGCCTCAAACTCGCTG 85558
Qy    1155 ct----atttatgatgattacgcacaccacccaacggaagtaactgcagtgctcagcgct 1210
Db   85559 GTGCGGGTGTTCGACGACTACGCACACCATCCAACGGAGATCAGCGCGACGCTAGCGGCG 85618
Qy    1211 gcgcgcacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacat 1270
Db   85619 TTCCGTATAATGCTTGAGCAGAGCGGCGGCGGACGCTGCATCGTGGTGTTTCAGCCCCAC 85678
Qy    1271 ttatactcacgcaccatagaattccaaaaggagttcgcgggggcactgtcactggcagac 1330
Db   85679 CTATATTCGCGGACAAAGGCTTTGTCCAGAGAATTTGGTCGTGCCCTGAGTGCTGCCGAT 85738
Qy    1331 gctgccgtggtgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcg 1390
Db   85739 GAGGTGTTCATCGTTGGTGTCAACGGTGCCCGCGAACAACCGCTTGCTGGCGTCAGCGGA 85798
Qy    1391 gaaatcatcaccgatgcgatgaccattccagtggtgtacgaacctaatttctctgcagtc 1450
Db   85799 GCTAGTGTTGCTAAACATGTCAGCGTGCCGGTGCGCTATATACCAGACTATTCGGCGGTC 85858
Qy    1451 ccagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggt 1510
Db   85859 GTCATAGAGGTTGCTGCTGCGGCTGGCCCCGGCGACGTCATCGTTACGATGGGTATCGGC 85918
Qy    1511 tccgtgaccatgcttgctccagaaatcctggatcagctgc 1550
Db   85919 GACGTGGGCCTGCTGGGGCCGGAGATCGTTGCCGCGCTGC 85958
```

```
RESULT 10
MSGB27CS
LOCUS       MSGB27CS    38793 bp    DNA            BCT       15-JUN-1996
DEFINITION  Mycobacterium leprae cosmid B27 DNA sequence.
ACCESSION   L78817
VERSION     L78817.1  GI:1377777
KEYWORDS    .
SOURCE      Mycobacterium leprae (clone: cosmid B27) (tissue library: Lorist 6)
            DNA.
  ORGANISM  Mycobacterium leprae
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (sites)
  AUTHORS   Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T.
  TITLE     Use of an ordered cosmid library to deduce the genomic organization
            of Mycobacterium leprae
  JOURNAL   Mol. Microbiol. 7 (2), 197-206 (1993)
  MEDLINE   93188700
REFERENCE   2  (bases 1 to 38793)
  AUTHORS   Smith,D.R., Richterich,P., Rubenfield,M., Butler,C., Lee,H.-M.,
            Xu,Q., Gunderson,K., Chung,M., Maher,J.K., Deloughery,C.,
            Aldrich,T., Imrich,J., Tulig,C., Smyth,A., Drill,S., Avruch,A.S.,
            Rice,P., Abendschan,K., Aldredge,T., DeLoughery,C., Kirst,S.,
            Safer,H., Connelly,S., McDougall,S., Eiglmeier,K., Bergh,S.,
            Cole,S., Robison,K., Jaehn,L., Gryan,G., Church,G.M. and Mao,J.
  TITLE     Prepublication submission
  JOURNAL   Unpublished (1996)
COMMENT     This sequence was generated by the Genome Sequencing Center at
            Genome Therapeutics Corporation (Collaborative Research Division),
            100
            Beaver St., Waltham, MA, 02154.  Please contact Doug Smith
            (smith@cric,com) for further information.  The sequence represents
            the
            insert of a Lorist 6 cosmid clone from a mapped set of clones
            constructed from M. leprae genomic DNA isolated from armadillo
            liver
            [3].  The sequence may not represent the entire cloned insert of
            the
            cosmid if an overlapping region was previously sequenced from
            another
            clone.  Coding sequences larger than 60 amino acids were predicted
            on
            the basis of codon usage and homology information.  An attempt was
            made
            to locate the most probable start site based on codon usage,
            homology,
            the presence of a Shine-Dalgarno sequence, or overlapping orf that
            suggested translational coupling.  It is possible that the actual
            start
            site differs from the one selected.
FEATURES             Location/Qualifiers
     source          1..38793
                     /organism="Mycobacterium leprae"
                     /specific_host="Dasypus novemcinctus"
                     /db_xref="taxon:1769"
                     /clone="cosmid B27"
                     /tissue_lib="Lorist 6"
                     /note="The liver of the armadillo was used to isolate the
                     Mycobacterium leprae."
BASE COUNT     7607 a  10325 c  11868 g   8993 t
ORIGIN

  Query Match             19.5%;  Score 308.4;  DB 1;  Length 38793;
  Best Local Similarity   53.6%;  Pred. No. 3.8e-64;
  Matches  757;  Conservative    0;  Mismatches  641;  Indels   14;  Gaps    5;

Qy     137 gatctgtcccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccga 196
Db    5765 GAGTTGCAGCGGGTGCACATGGTGGGCATTGGGGGATCCGGGATGTCGGGTATTGCGCGC 5824
Qy     197 atcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttg 256
```

```
Db  5825 ATCCTGTTAGACCGGGGCGGTCTGGTGTCTGGCTCAGATGCCAAGGAGTCGCGTGTTGTG 5884

Qy   257 cttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgag 316
Db  5885 CACGCACTGCGGGCGCGCGGCGCCTGATCCGCA-TCGGACATGATGCGTCATTGCTGGAC 5943

Qy   317 ctttccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaac 376
Db  5944 TTGCTGCCTGGCGGCGCTACCGCAGTCATTACCATCCGCACTGCCATCCCCAAGACTAAC 6003

Qy   377 ccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttg 436
Db  6004 CCCGAGCTGGTCGAAGCCCGGAGGCGCGGCATTCCGGTTCTATTGCGATCGGCCGTGTTG 6063

Qy   437 ggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctcc 496
Db  6064 GCCAGGCTGATGGATGGGTGTACCACGCTGATGGTTGCTGGCACACAGGCAAGACTACG 6123

Qy   497 accacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatc 556
Db  6124 ACGACGTCGATGCTCGTTGTGGCCCTGCAGCACTGTGGATGTGACCCATCCTTCGTGGTT 6183

Qy   557 ggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatc 616
Db  6184 GGCGGTGAGCTTGCCGTGGTGGGCACCAACGCCCACCACGGCCTCGGCGCCTGCTTCGTT 6243

Qy   617 gctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtc 676
Db  6244 GCCGAGGCCGACGAAAGCGATGGCTCGTTACTGGAATACACGCCCAACGTAGCGGTGGTC 6303

Qy   677 accaatgtggaaccagaccacctggacttctttaaaaccccctgaagcctacttccaagtg 736
Db  6304 ACCAACATCGACTCCGATCACTTGGATTTCTACGGTAGTGTCGATGCTTATATCAGGGTG 6363

Qy   737 ttcgacgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgat 796
Db  6364 TTCGACTCCTTCGTGGAGCGATTCGCACTGGGGGGAGCGCTCGTGGTGTGCAACGACGAT 6423

Qy   797 cctcacgcagcggagctgggggagaggtctgtccgcaaggtatcaagactgttggttac 856
Db  6424 CCTGGGGCGGCCGCGCTGGCCCGGCGCACCGCCGAGCTTGGGATCCGGGTGTTGCGCTAC 6483

Qy   857 ggtaccgctgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggat 916
Db  6484 GGCTCCGACGACCGGATAGGTGAGACCTTGGCCGCCAGGTTGTTGTCGT------GGGAG 6537

Qy   917 tcccaagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatct 976
Db  6538 CAACAGGGCACCGGAGCGGTCGCGCACATCCAGTTGGCTGGTCAACCAAACTCACGGGTG 6597

Qy   977 gtgattcttcaaatccctggtgatcacatggtactcaacggtgcagccgccctgctggcc 1036
Db  6598 ATGCGGCTACCGGTACCTGGGCGACACATGGCGCTCAATGCGCTAGGTGCGTTGCTGGCG 6657

Qy  1037 ggatacctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccggc 1096
Db  6658 GCGATCGAGATTGGTGCGTCGACCGATGAAGTGCTCGACGGGCTGGCTGGTTTCAGGGGT 6717

Qy  1097 gtgcgacgccgctttgagttccacggtgctatcgagggcggcaa--atttaatggcgctg 1154
Db  6718 GTGCGTCGCCGCTTCGAACTGGTTGGGACTTCAGGAGTCGGCCAGGCCTCAAACTCGCTG 6777

Qy  1155 ct----atttatgatgattacgcacaccacccaacggaagtaactgcagtgctcagcgct 1210
Db  6778 GTGCGGGTGTTCGACGACTACGCACACCATCCAACGGAGATCAGCGCGACGCTAGCGGCG 6837

Qy  1211 gcgcgcacccgggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacat 1270
Db  6838 TTCCGTATAATGCTTGAGCAGACGGCGGCGGACGCTGCATCGTGGTGTTTCAGCCCCAC 6897

Qy  1271 ttatactcacgcaccatagaattccaaaaggagttcgcgggggcactgtcactggcagac 1330

Db  6898 CTATATTCGCGGACAAAGGCTTTGTCCAGAGAATTTGGTCGTGCCCTGAGTGCTGCCGAT 6957

Qy  1331 gctgccgtggtgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcg 1390
Db  6958 GAGGTGTTCATCGTTGGTGTCAACGGTGCCCCGCGAACAACCGCTTGCTGGCGTCAGCGGA 7017

Qy  1391 gaaatcatcaccgatgcga-tgaccattccagtggtgtacgaacctaatttctctgcagt 1449
Db  7018 GCTAGTGTTGCTAAACAATGTCAGCGTGCCGGTGCGCTATATACCAGACTATTCGGCGGT 7077

Qy  1450 cccagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgcagg 1509
Db  7078 CGTCATAGAGGTTGCTGCTGCGGCTGGCCCCGGCGACGTCATCGTTACGATGGGTATCGG 7137

Qy  1510 ttccgtgaccatgcttgctccagaaatcctgg 1541
Db  7138 CGACGTGGGCCTGCTGGGGCCGGAGATCGTTG 7169
```

```
RESULT  11
AB003132
LOCUS       AB003132     4116 bp    DNA             BCT       04-AUG-1997
DEFINITION  Corynébacterium glutamicum gene for MurC, FtsQ, FtsZ, complete cds.
ACCESSION   AB003132
VERSION     AB003132.1  GI:2308989
KEYWORDS    FtsZ; FtsQ; MurC.
SOURCE      Corynebacterium glutamicum DNA.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 4116)
  AUTHORS   Kobayashi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-1997) to the DDBJ/EMBL/GenBank databases. Miki
            Kobayashi, Mitsubishi Chemical Corp., Tsukuba Research Center; Ami
            8-3-1, Inashiki, Ibaraki, Inashiki, Ibaraki 300-03, Japan
            (E-mail:3709292@cc.m-kagaku.co.jp, Tel:0298-87-1011,
            Fax:0298-87-3259)
REFERENCE   2  (sites)
  AUTHORS   Kobayashi,M., Asai,Y., Hatakeyama,K., Kijima,N., Wachi,M., Nagai,K.
            and Yukawa,H.
  TITLE     Cloning, sequencing, and characterization of the ftsZ gene from
            coryneform bacteria
  JOURNAL   Biochem. Biophys. Res. Commun. 236 (2), 383-388 (1997)
  MEDLINE   97382442
FEATURES             Location/Qualifiers
     source          1. .4116
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            1. .273
                     /gene="murC"
     CDS             <1. .273
                     /gene="murC"
                     /codon_start=1
                     /transl_table=11
                     /product="MurC"
                     /protein_id="BAA21685.1"
                     /db_xref="GI:2308990"
                     /translation="EFQKEPAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIP
                     VVYEPNFSAVPERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            301. .969
                     /gene="ftsQ"
     CDS             301. .969
                     /gene="ftsQ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsQ"
                     /protein_id="BAA21686.1"
                     /db_xref="GI:2308991"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKAQD
```

```
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKRDGQRWNISSPSM
                     VTVR"
     gene            1220. .2536
                     /gene="ftsZ"
     CDS             1220. .2536
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="FtsZ"
                     /protein_id="BAA21687.1"
                     /db_xref="GI:2308992"
                     /translation="MTSPNNYLAKIKVVGVGGGGDNAVNRMIEEGLKGVEFIAVNTDS
                     QALMFSDADVKLDIGREATRGLGAGANPEVGRASAETHKNEIEETIKGADMVFVTAGE
                     GGGAGTGAAPVVGRIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAALKEVCDTLIV
                     IPNDRLLELGDANLSIMEAFRAADEVLHNGVQGITNLITIPCINVDPADVRSVMSEAG
                     SALMGVGSARGDNRVVSATEQAINSPLLEATMDGATGVLLSFAGGSDLGLMEVNAAAS
                     MVRERSDEDVNLIFGTIIDDNLGDEVRVTVIATGFDAARASAAENRRAGIPAAPAAEP
                     VQQQQVPTTNATLPPEKESIFGGAREENDPYLSRSAGARHRIEETRSGGGLFTTGNDR
                     DYRRDERREDHRDERRDDRSYDRRDDRRDDRRDDRGDDLDVPSFLQ"
     CDS             2564. .3304
                     /note="hypothetical protein similar to YFIH"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA21688.1"
                     /db_xref="GI:2308993"
                     /translation="MDSLDPRNRPVRKVFTTRAGGVSQSPYATFNLGDHVGYYPKAVA
                     SNRNRLADIIGLSPDKVVYMEQIHSNTVTVIDEAPADGQAVEATDALVTTQRGLALAV
                     LVADCVPVLLSDTDAGVIAAVHAGRMGARNGIVAKTIAKMEELGAKPSRIHALMGAAA
                     SGAKYEVPEAMARDVEAKLPGSLARTTKGTTGLDIRAGLLRQMLSLGVQMIDSDPRCT
                     IEDEDLFSYRREGTTGRQAGVVWLPKEA"
     CDS             3345. .4010
                     /note="hypothetical protein similar to YPI1 or YPT5"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAA21689.1"
                     /db_xref="GI:2308994"
                     /translation="MQARIDATLNEHNRPEGSVRLLPVTKFHPVEDIKILQELGVTAV
                     GENREQEARAKALELPDMDFHMIGQIQSKKANSIARWSAAVHSVDSEKIAEALGRGVA
                     LALERGDRTSDELPCFIQLSLDGDPSRGGTPLSQVTQLADCISDTTHLRFEGLMCVPP
                     LGWDPEKAFSQARDVLSGLEEHFDRSLEFSAGNSGDLVAAIKHGSTIVRVGTEILGNR
                     PLA"
BASE COUNT     1009 a    1106 c    1148 g     853 t
ORIGIN


  Query Match          18.3%;  Score 289.8;  DB 1;  Length 4116;
  Best Local Similarity  99.3%;  Pred. No. 1.5e-59;
  Matches  291;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1289 gaattccaaaaggagttcgcgggggcactgtcactggcagacgctgccgtggtgcttgag 1348
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||| |||
Db      1 GAATTCCAAAAGGAGTTCGCGGAGGCACTGTCACTGGCAGACGCTGCCGTGGTGCTCGAG 60

Qy   1349 atttacggagcgcgcgaacaacggtggatggcgtgtcctcggaaatcatcaccgatgcg 1408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTTACGGAGCGCGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCG 120

Qy   1409 atgaccattccagtggtgtacgaaccttaatttctctgcagtcccagaacgcattgcagaa 1468
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATGACCATTCCAGTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAA 180

Qy   1469 atcgcaggacctaatgacatcgtgctcaccatgggtgcaggttccgtgaccatgcttgct 1528
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATCGCAGGACCTAATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCT 240

Qy   1529 ccagaaatcctggatcagctgcaaaacaattaggacgtaagtgaacaaggcag 1581
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCAGAAATCCTGGATCAGCTGCAAAACAATTAGGACGTAAGTGAACAAGGCAG 293


RESULT  12
BLFTSZ
LOCUS       BLFTSZ       5546 bp    DNA             BCT       19-JAN-2000
DEFINITION  Brevibacterium lactofermentum murC(partial), ftsQ, ftsZ genes
            ORF5, ORF6 DNA.
ACCESSION   Y08964
VERSION     Y08964.2  GI:6723445
KEYWORDS    cell division protein; divB gene; division initiation protein; ftsQ
            gene; ftsZ gene; murC gene; putative YAK 1 protein; putative YPT5
            protein; UDP-N-acetylmuramate-alanine ligase.
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 5546)
  AUTHORS   Honrubia,M.P., Fernandez,F.J. and Gil,J.A.
  TITLE     Identification, characterization, and chromosomal organization of
            the ftsZ gene from Brevibacterium lactofermentum
  JOURNAL   Mol. Gen. Genet. 259 (1), 97-104 (1998)
  MEDLINE   98409265
REFERENCE   2  (bases 1 to 5546)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1996) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 5546)
  AUTHORS   Gil,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-1999) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
COMMENT     On Jan 20, 2000 this sequence version replaced gi:1769958.
FEATURES             Location/Qualifiers
     source          1. .5546
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     gene            1. .264
                     /gene="murC"
     CDS             <1. .264
                     /gene="murC"
                     /EC_number="6.3.2.8"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-acetylmuramate--alanine ligase"
                     /protein_id="CAA70160.1"
                     /db_xref="GI:1769959"
                     /db_xref="SPTREMBL:P94335"
                     /translation="KEFAEALSLADAAVVLEIYGAREQPVDGVSSEIITDAMTIPVVY
                     EPNFSAVPERIAEIAGPNDIVLTMGAGSVTMLAPEILDQLQNN"
     gene            292. .960
                     /gene="ftsQ"
     CDS             292. .960
                     /gene="ftsQ"
                     /function="division initiation protein or cell division
                     protein"
                     /note="alternative gene name divB"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70161.1"
                     /db_xref="GI:1769960"
                     /db_xref="SPTREMBL:P94336"
                     /translation="MNKKVIAIVVGVVVVLVAILGVVAWFVPILKVGNIEVTGATRTD
                     PDQVLEVSGIVEGENLFRVDATAAGQNIVELPWVKSVTVNRALPSTITVELTEREPAV
                     FIKRADGDHVIDTEGKEIIIGTPPVGTVEVSGADEGNSEVLPAVIAVINAIKRQDAQM
                     TESIQVVEAPDQFDILLKMNDGREIYWGSSENNHDKAVAMSTVLKREGQRWNISSPSM
                     VTVR"
     gene            1211. .2527
                     /gene="ftsZ"
     CDS             1211. .2527
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein"
                     /protein_id="CAA70158.1"
                     /db_xref="GI:1769961"
```

```
                     /db_xref="SWISS-PROT:P94337"
                     /translation="MTSPNNYLAKIKVVGVGGGCVNAVNRMIEEGLKGVEFIAVNTDS
                     QALMFSDADVKLDIGREATRGLGAGANPEVGRASAEDHKNEIEETIKGADMVFVTAGE
                     GGGTGTGAAPVVAGIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAALKEVCDTLIV
                     IPNDRLLELGDANLSIMERFRAADEVLHNGVQGITNLITIPGVINVDFADVRSVMSEA
                     GSALMGVGSARGDNRVVSATEQAINSPLLEATMDGATGVLLSFAGGSDLGLMEVNAAA
                     SMVRERSDEDVNLIFGTIIDDNLGDEVRVTVIATGFDAARASAAENRRAGIPAAPAAE
                     PVQQQVPTTNATLPPEKESIFGGAREENDPYLSRSAGARHRIEETRSGGGLFTTGNDR
                     DYRRDERREDHRDERRDDRSYDRRDDRRDDRRDDRGDDLDVPSFLQ"
     gene            2555. .3295
                     /gene="yfih"
     CDS             2555. .3295
                     /gene="yfih"
                     /note="hypothetical protein similar to YFIH protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70159.1"
                     /db_xref="GI:1769962"
                     /db_xref="SPTREMBL:P94338"
                     /translation="MDSLDPRNRPVRKVFTTRAGGVSQSPYASFNLGDHVGDDPQAVA
                     SNRNRLANIIGLSPDKVVYMEQIHSNTITVIDEAPADGQAVEATDALVTTQRGLALAV
                     LVADCVPVLLSDTDAGVIAAVHAGRMGARNGIVAKTIAKMEELGAKPSRIHALMGAAA
                     SGANYEVPEAMARDVEAKLPGSIARTTKGTTGLDIRAGLLRQMLSLGVQMIDSDPRCT
                     IEDEDLFSYRREGTTGRQAGVVWLPKEA"
     CDS             3376. .4041
                     /note="ORF5"
                     /codon_start=1
                     /transl_table=11
                     /product="putative YPT5 protein"
                     /protein_id="CAB66326.1"
                     /db_xref="GI:6723446"
                     /translation="MQARIDATLNEHNRPEGSVRLLPVTKFHPVEDIKILQELGVTAV
                     GENREQEARAKALELPDMDFHMIGQIQSKKANSIARWAAAVHSVDSEKIAEALGRGVA
                     LALERGDRTGDELPCFIQLSLDGDPSRGGTPLSQVTQLADCISDTTHLRFEGLMCVPP
                     LGWDPEKAPSQARDVLSGLEEEHFDRSLEFSAGMSGDLIAAIKHGSTIVRVGTEILGNR
                     PLA"
     CDS             4118. .4576
                     /note="ORF6"
                     /codon_start=1
                     /transl_table=11
                     /product="putative YAK 1 protein"
                     /protein_id="CAB66327.1"
                     /db_xref="GI:6723447"
                     /translation="MSMLEKTKEFPGLAPYEAEHEDAYYADEPRYEGTAAYAPEYRER
                     DIGYAPEAPAPVAPSPAPRSIQSTIVPVELHSFEDAQVIGGAFRDGDAVVFDMSLLSR
                     EEARRIVDFAAGLSFALRGKMQKIDSVTFAVVPELSNISTSELERAARIR"
BASE COUNT     1317 a   1479 c   1520 g   1230 t
ORIGIN

  Query Match          17.8%;  Score 280.8;  DB 1;  Length 5546;
  Best Local Similarity   99.3%;  Pred. No. 2.2e-57;
  Matches  282;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   1298 aaggagttcgcgggggcactgtcactggcagacgctgccgtggtgcttgagatttacgga 1357
          ||||||||||||| ||||||||||||||||||||||||||||||||||| ||||||||||||
Db      1 AAGGAGTTCGCGGAGGCACTGTCACTGGCAGACGCTGCCGTGGTGCTCGAGATTTACGGA 60

Qy   1358 gcgcgcgaacaaccggtggatggcgtgtcctcggaaatcatcaccgatgcgatgaccatt 1417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCGCGCGAACAACCGGTGGATGGCGTGTCCTCGGAAATCATCACCGATGCGATGACCATT 120

Qy   1418 ccagtggtgtacgaacctaatttctctgcagtcccagaacgcattgcagaaatcgcagga 1477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCAGTGGTGTACGAACCTAATTTCTCTGCAGTCCCAGAACGCATTGCAGAAATCGCAGGA 180

Qy   1478 cctaatgacatcgtgctcaccatgggtgcaggttccgtgaccatgcttgctccagaaatc 1537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTAATGACATCGTGCTCACCATGGGTGCAGGTTCCGTGACCATGCTTGCTCCAGAAATC 240

Qy   1538 ctggatcagctgcaaaacaattaggacgtaagtgaacaaggcag 1581
          |||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGATCAGCTGCAAAACAATTAGGACGTAAGTGAACAAGGCAG 284
```

```
RESULT  13
SC9B1
LOCUS       SC9B1       24800 bp    DNA             BCT       27-APR-1999
DEFINITION  Streptomyces coelicolor cosmid 9B1.
ACCESSION   AL049727
VERSION     AL049727.1  GI:4691379
KEYWORDS    A-factor receptor homolog; ABC transporter; ATP/GTP-binding;
            carbonic anhydrase; cprB; cyclase; dipeptidase; hydrolase; integral
            membrane; oxidoreductase; permease; secreted substrate-binding;
            tetR; transferase; UDP-N-acetylmuramoyl-L-alanine ligase.
SOURCE      Streptomyces coelicolor A3(2).
  ORGANISM  Streptomyces coelicolor A3(2)
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 24800)
  AUTHORS   Redenbach,M., Kieser,H.M., Denapaite,D., Eichner,A., Cullum,J.,
            Kinashi,H. and Hopwood,D.A.
  TITLE     A set of ordered cosmids and a detailed genetic and physical map
            for the 8 Mb Streptomyces coelicolor A3(2) chromosome
  JOURNAL   Mol. Microbiol. 21 (1), 77-96 (1996)
  MEDLINE   97000351
REFERENCE   2  (bases 1 to 24800)
  AUTHORS   Saunders,D.C. and Harris,D.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 24800)
  AUTHORS   Bentley,S.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-1999) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT     Notes:
            Streptomyces coelicolor sequencing at The Sanger Centre is funded
            by the BBSRC.
            Details of S. coelicolor sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL; http://www.sanger.ac.uk/Projects/S_coelicolor/)
            CDS are numbered using the following system eg SC7B7.01c. SC (S.
            coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous.
            The length in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
            CDS which show significant similarity to other CDS in the database.
            The position of possible ribosome binding site sequences are given
            where these have been used to deduce the initiation codon.
            Gene prediction is based on positional base preference in codons
            using a specially developed Hidden Markov Model (Krogh et al.,
            Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot
            program of Bibb et al., Gene 30:157-66(1984) as implemented at
            http://www.nih.go.jp/
            jun/cgi-bin/frameplot.pl. CAUTION:  We may not have predicted the
            correct initiation codon.  Where possible we choose an initiation
            codon (atg, gtg, ttg or (att)) which is preceded by an upstream
            ribosome binding site sequence (optimally 5-13bp before the
            initiation codon).  If this cannot be identified we choose the most
            upstream initiation codon.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone.  It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
            small overlap between neighbouring submissions.
            Cosmid 9B1 lies between 1B5 and 4A1 on the AseI-A genomic
            restriction fragment.
FEATURES             Location/Qualifiers
     source          1. .24800
                     /organism="Streptomyces coelicolor A3(2)"
                     /strain="A3(2)"
                     /db_xref="taxon:100226"
                     /clone="cosmid 9B1"
```

```
gene            complement(1. .447)
                /gene="SC9B1.01c"
misc_feature    1. .103
                /note="Nominal overlap with Streptomyces coelicolor cosmid
                1B5."
CDS             complement(<1. .447)
                /gene="SC9B1.01c"
                /note="SC9B1.01c, partial CDS, possible integral membrane
                protein, len: >149aa; similar to many eg. TR:CAB38499
                (EMBL:AL035636) putative integral membrane protein from
                Streptomyces coelicolor (830 aa) fasta scores; opt: 385,
                z-score: 428.6, E(): 1.5e-16, (44.7% identity in 141 aa
                overlap). Contains several possible membrane spanning
                hydrophobic regions."
                /codon_start=1
                /transl_table=11
                /label=SC9B1.01c
                /product="putative integral membrane protein"
                /protein_id="CAB41547.1"
                /db_xref="GI:4691380"
                /translation="MSPIADRNRRFPHWRQDLTASLVVFLVALPLCVGVAVASGVPAE
                LGLVTGIVGGLVTGFLPGSSLQVSGPAAGLTVLVFGAVDQYGLPALGVIVLATGIVQL
                AMGALKLGRWFRAISLSVVEGMLAGIGLVLIAGQLYSAAGLQAPASG"
gene            complement(456. .1037)
                /gene="SC9B1.02c"
CDS             complement(456. .1037)
                /gene="SC9B1.02c"
                /note="SC9B1.02c, probable carbonic anhydrase, len: 193aa;
                similar to many eg. SW:CYNT_ECOLI carbonic anhydrase from
                Escherichia coli (219 aa) fasta scores; opt: 512, z-score:
                618.8, E(): 3.8e-27, (43.6% identity in 188 aa overlap).
                Contains two Pfam matches to entry PF00484 Pro_CA,
                Prokaryotic-type carbonic anhydrases and Prosite matches
                to PS00705 Prokaryotic-type carbonic anhydrases signature
                2 and PS00704 Prokaryotic-type carbonic anhydrases
                signature 1."
                /codon_start=1
                /transl_table=11
                /label=SC9B1.02c
                /product="probable carbonic anhydrase"
                /protein_id="CAB41548.1"
                /db_xref="GI:4691381"
                /translation="MQPLIDNAMFGQRPEEFTGHAEGQSPEVLFITCSDSRVVPALI
                TGARPGQLFELRTAGNIVPPHGTGHPSGEAATVEYAVQVLGVADIVVCGHSHCGAVGA
                LVRGDDLTAVPAVRDWLAHAADEPKPCDPADPTVAGAVQHHVLAQLLRLRSYPCVERR
                LADGRLRLHGWYYEVHTGLVRAHRADTDAFETL"
misc_feature    complement(504. .608)
                /gene="SC9B1.02c"
                /note="Pfam match to entry PF00484 Pro_CA,
                Prokaryotic-type carbonic anhydrases, score 31.50, E-value
                4.7e-08."
misc_feature    complement(666. .1019)
                /gene="SC9B1.02c"
                /note="Pfam match to entry PF00484 Pro_CA,
                Prokaryotic-type carbonic anhydrases, score 145.30,
                E-value 4.5e-41."
misc_feature    complement(744. .806)
                /gene="SC9B1.02c"
                /note="PS00705 Prokaryotic-type carbonic anhydrases
                signature 2."
misc_feature    complement(915. .938)
                /gene="SC9B1.02c"
                /note="PS00704 Prokaryotic-type carbonic anhydrases
                signature 1."
gene            complement(1116. .1517)
                /gene="SC9B1.03c"
CDS             complement(1116. .1517)
                /gene="SC9B1.03c"
                /note="SC9B1.03c, possible secreted protein, len: 133aa;
                contains possible N-terminal signal sequence."
                /codon_start=1
                /transl_table=11
                /label=SC9B1.03c
                /product="putative secreted protein"
                /protein_id="CAB41549.1"
                /db_xref="GI:4691382"
                /translation="MYRRTAARTVVAVLAPVLFALQFFAPSGSFASAYTNGDAVAHAR
                PGTIPSGTAPHDETVTCREGGRPGEPTTAPGVRDRHRTTAAPQSDTPRHPPARQSAPQ
                APGSVASGRATGHPPRAAADHAPAVLQVFRC"
gene            complement(1707. .2759)
                /gene="SC9B1.04c"
CDS             complement(1707. .2759)
                /gene="SC9B1.04c"
                /note="SC9B1.04c, possible ATP/GTP-binding integral
                membrane protein, len: 350aa; similar to hypotheticals eg.
                TR:P71967 (EMBL:Z80225) from Mycobacterium tuberculosis
                (369 aa) fasta scores; opt: 1129, z-score: 1182.3, E(): 0,
                (57.8% identity in 325 aa overlap). Contains possible
                membrane spanning hydrophobic regions. Also contains
                Prosite match to PS00017 ATP/GTP-binding site motif A
                (P-loop)."
                /codon_start=1
                /transl_table=11
                /label=SC9B1.04c
                /product="putative ATP/GTP-binding integral membrane
                protein"
                /protein_id="CAB41550.1"
                /db_xref="GI:4691383"
                /translation="MSLCAREPRVPADRLVAEMVPPPRFDSVRFGTYVPDPNQPSQSE
                AVRVLEDFAGGLGEAPGTGGGRRGFFGLGRGRAPKTPAGPRGVYLDGGYGVGKTHLLA
                SLWHATPAEPSRKAFGTFVELTNLVGALGFQQTVQTLSGHRLLCIDEFELDDPGDTVL
                VSSLLARLVEAGVALAATSNTLPGKLGEGRFAAADFLREIQGLSAHFRALRIDGEDYR
                HRGLPEAPAPYSDEQVTRAARATEGASLDDPAALLDHLARVHPSRYGALTDGLTAVCL
                TGVHPVPDQSTALRLVVLADRLYDREVPVLASGLPFDRLFGEEMLKGGYRKKYFRAIS
                RLTALARDAARLVDTA"
misc_feature    complement(2466. .2489)
                /gene="SC9B1.04c"
                /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
repeat_region   2863. .3232
                /note="High GC content."
gene            3340. .4137
                /gene="SC9B1.05"
CDS             3340. .4137
                /gene="SC9B1.05"
                /note="SC9B1.05, possible hydrolase, len: 265aa; similar
                to hypotheticals eg. TR:P71968 (EMBL:Z80225) from
                Mycobacterium tuberculosis (258 aa) fasta scores; opt:
                536, z-score: 585.9, E(): 2.6e-25, (44.3% identity in 235
                aa overlap). Also similar to several proposed ribofavin
                deaminases eg. SW:RIBG_BACAM ribofavin-specific deaminase
                from Bacillus subtilis (371 aa) fasta scores; opt: 157,
                z-score: 177.0, E(): 0.015, (28.3% identity in 180 aa
                overlap)."
                /codon_start=1
                /transl_table=11
                /label=SC9B1.05
                /product="putative hydrolase"
                /protein_id="CAB41551.1"
                /db_xref="GI:4691384"
                /translation="MRRLFPVSPATDPTAGPGGGPETDREWSLAELAAAYAYPEPDAH
                RRTWLRANMVSTLDGAAQHDGRSQPISSAADMRVFGTLRALADVVIAGAETVRQEGYR
```

```
Query Match          14.7%;  Score 231.8;  DB 1;  Length 24800;
Best Local Similarity   51.8%;  Pred. No. 1.3e-45;
Matches  747;  Conservative   .0;  Mismatches  622;  Indels   72;  Gaps    7;

Qy    141 tgtcccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccgaatcc 200
          ||  |||    ||| ||||||| || ||||| |||||||||||| ||| | |||   ||||
Db   4926 TGGACCGACCGCACTTCATCGGCATCGGCGGCGCCGGAATGTCGGGCATCGCCAAGATCC 4985

Qy    201 tgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttgcttc 260
          |      |||||       ||  | ||   ||| |||||  || |||   |||
Db   4986 TCGCCCAGCGCGGCGCCGAGGTGGCGGGCAGCGACGCCAAGGAGTCCGCGACCGCCGACG 5045

Qy    261 cactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgagcttt 320
          | || |||||| |||| |||||| |     | || ||||| |||||  |||| |
```

```
Db   5046 CGCTGCGCGCCCTGGGCGCCACCGTGCACATCGGGCACGCGGCGGAGCACCTCG------ 5099

Qy    321 ccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaacccgg 380
Db   5100 CCGCCGACGCGAGCTGCGTGGTCGTGTCCT------CGGCCATCCGCGAGGACAATCCGG 5153

Qy    381 aacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttgggcg 440
Db   5154 AGCTGGTCCGCGCCGCCGAGCTGGGCATCCCGGTCGTGCACCGCTCCGACGCCCTCGCCG 5213

Qy    441 aattgctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctccacca 500
Db   5214 CCCTGATGAACGGCCTGCGCCCCATCGCGGTCGCCGGCACCCACGGCAAGACCACCACCA 5273

Qy    501 cctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatcggcg 560
Db   5274 CCTCGATGCTGGCGGTCTCCCTGTCCGAGCTGGACCTGGGCCCGTCGTACGCGATCGGCG 5333

Qy    561 gacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatcgctg 620
Db   5334 GCGACCTGGACGCCCCCGGCTCCAACGCCCTGCACGGCGAAGGCGAGATCTTCGTCGCCG 5393

Qy    621 aagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtcacca 680
Db   5394 AGGCGGACGAAAGCGACCGCAGCTTCCACAAGTACGCTCCCGAGGTCGCCATCGTCCTCA 5453

Qy    681 atgtggaaccagaccacctggacttctttaaaacccctgaagcctacttccaagtgttcg 740
Db   5454 ACGTCGAGCTGGACCACCACGCCAACTACGCCTCGATGGACGAGATCTACGAGTCCTTCG 5513

Qy    741 acgatttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgatcctc 800
Db   5514 AGACCTTCGCCGGCAAGATCGTCCCCGGCGGCACCCTGGT------------GATCGCG 5560

Qy    801 acgcagcggagctggggagaggtctgtccgcaagggtatcaagactgttggttacggta 860
Db   5561 GCCGACCACGAGGGCGCCCGGGAGCTGACCCGACGGCTGGCAGGCCGGGTGCGGACGGTG 5620

Qy    861 ccgctgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggattccc 920
Db   5621 ACGTACGGAGAGTCCGAGGACGCCGACGTGCGCATCCTGTCCGTCGTCCCGCAGGGCCTC 5680

Qy    921 aagttgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatctgtga 980
Db   5681 AAG-------------AGCGAGGTCACCGTCGTCCTGGACGGCGCCGAGCTGACCTTCG 5726

Qy    981 ttcttcaaatccctggtgatcacatggtactcaacggtgcagccgccctgctggccggat 1040
Db   5727 CCGTCTCCGTCCCCGGCCGCCACTACGCCCACAACGCCGTCGCCGCCCTCGCCGCGGGCG 5786

Qy   1041 acctggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccggcgtgc 1100
Db   5787 CCGCCCTCGGCGTCCCCGCCGCCGAGCTGGCCCCCGCGCTGGCCGCCTACACGGGCGTCA 5846

Qy   1101 gacgccgctttgagttccacggtgctatcgagggcggcaaatttaatggcgctgctattt 1160
Db   5847 AGCGGCGCCTCCAGCTCAAGGGCGAGGCGGCCGGCGTCCAGGTCGTCGACTC--------- 5898

Qy   1161 atgatgattacgcacaccacccaacggaagtaactgcagtgctcagcgctgcgcgcaccc 1220
Db   5899 -------CTACGCGCACCACCCCACCGAGATGACCGCCGACCTGGAGGCCATGCGCGCC- 5950

Qy   1221 gggtgaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacatttatactcac 1280
Db   5951 --------GCGGTCGGCGACGCCCGCATCCTGGTCCTCTTCCAGCCGCACCTCTTCTCCC 6002

Qy   1281 gcaccatagaattccaaaaggagttcgcggggcactgtcactggcagacgctgccgtgg 1340
Db   6003 GCACCCAGGAGCTGGGCAAGGAGATGGGCCAGGCCCTGGCCCTCGCGGACGCCTCGGTCG 6062

Qy   1341 tgcttgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcggaaatcatca 1400
Db   6063 TCCTCGACATCTACCCGGCCCGCGAGGACCCGATCCCGGGCGTCACCAGCGAGCTGATCG 6122

Qy   1401 ccgatgcgatgaccattccagtggtgtacgaacctaatttc---------tctgcagtcc 1451
Db   6123 TCGAGGCGGCCCGCGCCGCCGGCGCCGACGTCACACCGGTCCACGACAAGGACGCCTCGC 6182

Qy   1452 cagaacgcattgcagaaatcgcaggacctaatgacatcgtgctcaccatgggtgcaggtt 1511
Db   6183 CCGCGCTGGTCGCGGGAATGGCGAAGGCCGGTGATCTCGTTCTCACCATGGGCGCGGGCG 6242

Qy   1512 ccgtgaccatgcttgctccagaaatcctggatcagctgcaaaacaattaggacgtaagtg 1571
Db   6243 ATGTCACCGACCTCGGACCGCGCATCCTGGACGAGCTGTCGAAGTAAGGGGCTGAGACTC 6302

Qy   1572 a 1572
Db   6303 A 6303
```

```
RESULT 14
MLCB268/c
LOCUS       MLCB268     38859 bp    DNA             BCT       27-AUG-1999
DEFINITION  Mycobacterium leprae cosmid B268.
ACCESSION   AL022602
VERSION     AL022602.1  GI:3080471
KEYWORDS    3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; anthranilate
            phosphoribosyltransferase; aroG; ctaE; cytochrome c oxidase subunit
            III; ftsW; long chain fatty acid-coA ligase; murC; murD; murE;
            murF; murG; murX; pbpB; penicillin binding protein 2;
            phospho-n-acetylmuramoyl-pentapeptide-transferase; pseudogene;
            serine/threonine protein kinase; trpD; udp-n-acetylmuramate-alanine
            ligase; udp-n-acetylmuramoylalanine--d-glutamate ligase;
            udp-n-acetylmuramoylalanyl-d-glutamate--2,6-diaminopimelate ligase.
SOURCE      Mycobacterium leprae.
  ORGANISM  Mycobacterium leprae
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (bases 1 to 38859)
  AUTHORS   Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T.
  TITLE     Use of an ordered cosmid library to deduce the genomic organization
            of Mycobacterium leprae
  JOURNAL   Mol. Microbiol. 7 (2), 197-206 (1993)
  MEDLINE   93188700
REFERENCE   2  (bases 1 to 38859)
  AUTHORS   Brown,D. and Churcher,C.M.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 38859)
  AUTHORS   Parkhill,J., Barrell,B.G. and Rajandream,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-1997) Mycobacterium leprae sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Dr.
            Stewart T. Cole, [3] Unite de Genetique Moleculaire Bacterienne,
            Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15,
            France Requests for cosmids should be sent to Karin Eiglmeier
            (kei@pasteur.fr)
COMMENT     Notes:
            The Sanger Centre is funded to complete the sequence of M. leprae
            by the Heiser Program for Research in Leprosy and Tuberculosis of
            The New York Community Trust.
            Work in Paris is supported by the Heiser Trust, the Association
            Francaise Raoul Follereau and the Groupement de Recherches et des
            Etudes des Genomes (GIP-GREG).
            Details of M. leprae sequencing at the Sanger Centre are available
            on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/)
            CDS are numbered using the following system eg MLCB33.01c. ML (M.
            leprae), cB33 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous. The length
            in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
```

CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. All CDS over 100 codons have been analysed. Gene prediction is based on positional base preference in codons especially where there is an increase in the observed/expected third position G + C. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid B268 is overlapped at the 5' end by MLB1554CS and at the 3' end by MLB1551CS.

FEATURES Location/Qualifiers

source 1. .38859
/organism="Mycobacterium leprae"
/db_xref="taxon:1769"
/clone="cosmid B268"

gene complement(1. .705)
/gene="murC"

CDS complement(<1. .705)
/gene="murC"
/note="MLCB268.01c, murC, probable udp-n-acetylmuramate-alanine ligase, partial CDS, len: >231 aa; highly similar to many e.g. MURC_ECOLI P17952 udp-n-acetylmuramate-alanine ligase (EC 6.3.2.8) (491 aa), fasta scores; opt: 542 z-score:704.5 E(): 4.1e-32, 41.9% identity in 227 aa overlap"
/codon_start=1
/transl_table=11
/product="udp-n-acetylmuramate-alanine ligase"
/protein_id="CAA18667.1"
/db_xref="GI:3080472"
/db_xref="SPTREMBL:O69551"
/translation="MNAGQLPPELQRVHMVGIGGSGMSGIARILLDRGGLVSGSDAKESRVVHALRARGALIRIGHDASLLDLLPGGATAVITIRTAIPRTNPELVEARRRGIPVLLRSAVLARLMDQCTTLMVAGTHGKTTTTSMLVVALQHCGCDPSFVVGGELAVVGTNAHHGLGACFVAEADESDGSLLEYTPNVAVVTNIDSDHLDFYGSVDAYIRVFDSFVERFALGGALVVCNDDPGAAALA"

misc_feature complement(1. .6446)
/note="overlap with MLB27CS from: 1 to: 6444. There are sixteen conflicts between this sequence and ours. In each case our sequence has been checked, and is thought to be correct"

misc_feature 538
/note="conflict: this base is missing in MLB27CS"

gene complement(702. .1925)
/gene="murG"

CDS complement(702. .1925)
/gene="murG"
/note="MLCB268.02c, murG, probable udp-n-acetylglucosamine--n-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol n-acetylglucosamine transferase, len: 407 aa; highly similar to many e.g. MURG_ECOLI P17443 (354 aa), fasta scores; opt: 461 z-score: 727.7 E(): 0, 38.1% identity in365 aa overlap"
/codon_start=1
/transl_table=11
/product="udp-n-acetylglucosamine--n-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol n-acetylglucosamine transferase"
/protein_id="CAA18668.1"
/db_xref="GI:3080473"
/db_xref="SPTREMBL:O69552"
/translation="MNNSVREPTRGRRGSPPVADAALSVNPPLSVVLAGGGTAGHVEPAMAVADALRALDPQVRITALGTSRGLETRLVPERGYHLELITPVPLPRKLTGDLARLPLRVWRAVRETRAVFEVVEAHVVVGPGGYVALPAYLAARGIPRVRRRIPVVVHEANARAGIANRVGVRTAERVLSAVPGSGLRGAEVVGVPIHATITTLNRPALRADARKHFGFTDDARVLLVFGGSQGAVSLNRAVAGAAEDLAASGVAVLHAYGLRNTLELRTPEYGEPPYVAVPYLDRMDLAYAAADLVICRSGAMTVAEVSAVGLPAIYVPPPIGNGEQRLNALPVVNAGGGLVVADADLTPGLVARQVVRLFSDPAQLAAMTAAAARVGHRDAAHHVAKVALDLARAERDTASGRSAGGKP"

RBS complement(712. .716)
/gene="murG"
/note="possible RBS upstream of murC"

misc_feature 1320
/note="conflict: G is C in MLB27CS"

misc_feature 1819
/note="conflict: C is T in MLB27CS"

misc_feature 1824
/note="conflict: this base is missing in MLB27CS"

gene complement(1922. .3526)
/gene="ftsW"

CDS complement(1922. .3526)
/gene="ftsW"
/note="MLCB268.03c, ftsW, probable cell division protein, len: 534 aa; similar to many e.g. FTSW_ECOLI P16457 cell division protein ftsW (414 aa), fasta scores; opt: 647 z-score: 573.9 E(): 7.7e-28, 37.0% identity in 370 aa overlap. Contains PS00428 Cell cycle proteins ftsW / rodA / spoVE signature and PS00041 Bacterial regulatory proteins, araC family signature"
/codon_start=1
/transl_table=11
/product="cell divisin protein FtsW"
/protein_id="CAA18669.1"
/db_xref="GI:3080474"
/db_xref="SPTREMBL:O69553"
/translation="MGNALGRDRGRGTGDNNSVQAATTVEAVEPGTATTGEVSEARAKRAGSLRAGVRGPQTRFDAWLGRPMASFHLIIAVAGLLTALGLIMVLSASGVRSYGDDGSAWIIFGKQVLWTIIGLIGGTSSLWMSIRFIRRIAFFSYVITIILLVLVLIPGIGNLANGSRKWFVIAGFSMQPSELAKIAFVIWGAHLLAARRMERASLREMLIPLVPAAVIALGLIVAQPDLGQTVSLGIILLALLWYAGLPLRVFITSLLAVFIAGAILAMSAGYRSERVRSWMNPEADPQDTGYQARQAKFALAHGGIFGDGLGQGVAKWNYLPNAHNDFIPAIIGEELGLVGALGLLGLFGLFAYTGMRIARRSADPFLRLLTATTTMWVLGQAFINIGTVIGVLPVTGLQLPFISAGGTSAAAILFMIGIMANAARHEPEAVAALRAGRDDKVNQLLRLPLPVPYLPPRVETFRDRKQVRPQPGKASVGRKPARKAPREPDRPLRPAPPRVADRLEHRSGQYRAGQLRAGQRYTPRVRVPEGQRYG"

RBS complement(1937. .1942)
/gene="ftsW"
/note="possible RBS upstream of murG"

misc_feature 2019. .2020
/note="conflict: AG is AAG in MLB27CS"

misc_feature 2053
/note="conflict: this base is missing in MLB27CS"

misc_feature 2061. .2062
/note="conflict: TC is CA in MLB27CS"

misc_feature 2081
/note="conflict: this base is missing in MLB27CS"

misc_feature 2163
/note="conflict: G is T in MLB27CS"

misc_feature complement(2300. .2428)
/gene="ftsW"
/note="PS00041 Bacterial regulatory proteins, araC family signature"

misc_feature complement(2303. .2377)
/gene="ftsW"
/note="PS00428 Cell cycle proteins ftsW / rodA / spoVE signature"

misc_feature 2312. .2313
/note="conflict: AC is AACC in MLB27CS"

misc_feature 2973. .2974
/note="conflict: AC is AAC in MLB27CS"

misc_feature 2989
/note="conflict: this base is missing in MLB27CS"

RBS complement(3529. .3534)
/note="possible RBS upstream of ftsW"

gene complement(3533. .5005)
/gene="murD"

CDS complement(3533. .5005)
/gene="murD"
/note="MLCB268.04c, murD, probable

udp-n-acetylmuramoylalanine--d-glutamate ligase, len: 490

Query Match 14.3%; Score 226.2; DB 1; Length 38859;
Best Local Similarity 58.3%; Pred. No. 2.8e-44;
Matches 396; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

```
Qy    137 gatctgtcccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccga 196
Db    681 GAGTTGCAGCGGGTGCACATGGTGGGCATTGGGGGATCCGGGATGTCGGGTATTGCGCGC 622

Qy    197 atcctgcttgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttg 256
Db    621 ATCCTGTTAGACCGGGGCGGTCTGGTGTCTGGCTCAGATGCCAAGGAGTCGCGTGTTGTG 562

Qy    257 cttccactccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaaccttgag 316
Db    561 CACGCACTGCGGGCGCGCGGCGCGCTGATCCGCATCGGACATGATGCGTCATTGCTGGAC 502

Qy    317 ctttccggcgaacttcccaccgtcgtggtgacctcttttgccgccattccgcaagacaac 376
Db    501 TTGCTGCCTGGCGGCGCTACCGCAGTCATTACCATCCGCACTGCCATCCCCAAGACTAAC 442

Qy    377 ccggaacttgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttg 436
Db    441 CCCGAGCTGGTCGAAGCCCGGAGGCGCGGCATTCCGGTTCTATTGCGATCGGCCGTGTTG 382

Qy    437 ggcgaattgctggaaggctccacccaggtcttgatcgcgggtacccacgggtaagacctcc 496
Db    381 GCCAGGCTGATGGATGGGTGTACCACGCTGATGGTTGCTGGCACACACGGCAAGACTACG 322

Qy    497 accacctctatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatc 556
Db    321 ACGACGTCGATGCTCGTTGTGGCCCTGCAGCACTGTGGATGTGACCCATCCTTCGTGGTT 262

Qy    557 ggcggacagctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatc 616
Db    261 GGCGGTGAGCTTGCCGTGGTGGGCACCAACGCCCACCACGGCCTCGGCGCCTGCTTCGTT 202

Qy    617 gctgaagcagatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtc 676
Db    201 GCCGAGGCCGACGAAAGCGATGGCTCGTTACTGGAATACACGCCCAACGTAGCGGTGGTC 142

Qy    677 accaatgtggaaccagaccacctggacttctttaaaaacccctgaagcctacttccaagtg 736
Db    141 ACCAACATCGACTCCGATCACTTGGATTTCTACGGTAGTGTCGATGCTTATATCAGGGTG 82

Qy    737 ttcgacgatttcgcaggacgcatcaccccgaacggcaaggctggttgtgtgcctgaacgat 796
Db     81 TTCGACTCCTTCGTGGAGCGATTCGCACTGGGGGGAGCGCTCGTGGTGTGCAACGACGAT 22

Qy    797 cctcacgcagcggagctgg 815
Db     21 CCTGGGGCGGCCGCGCTGG 3
```

RESULT 15
AP002997/c
LOCUS AP002997 329709 bp DNA BCT 15-MAY-2001
DEFINITION Mesorhizobium loti DNA, complete genome, section 4/21.
ACCESSION AP002997 BA000012
VERSION AP002997.2 GI:14022051
KEYWORDS
SOURCE Mesorhizobium loti (strain:MAFF303099) DNA.
ORGANISM Mesorhizobium loti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Phyllobacteriaceae; Mesorhizobium.
REFERENCE 1 (sites)
AUTHORS Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S., Watanabe,A., Idesawa,K., Ishikawa,A., Kawashima,K., Kimura,T., Kishida,Y., Kiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A., Mochizuki,Y., Nakayama,S., Nakazaki,N., Shimpo,S., Sugimoto,M., Takeuchi,C., Yamada,M. and Tabata,S.
TITLE Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti
JOURNAL DNA Res. 7 (6), 331-338 (2000)
MEDLINE 21082930
REFERENCE 2 (bases 1 to 329709)
AUTHORS Kaneko,T.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail:kaneko@kazusa.or.jp,
URL:http://www.kazusa.or.jp/rhizobase/,
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
COMMENT On May 11, 2001 this sequence version replaced gi:11994965.
FEATURES Location/Qualifiers
source 1. .329709
/organism="Mesorhizobium loti"
/strain="MAFF303099"
/db_xref="taxon:381"
gene complement(79. .402)
/gene="mll1244"
CDS complement(79. .402)
/gene="mll1244"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48662.1"
/db_xref="GI:14022052"
/translation="MSRKGLSIPPNGSPHYWADCLMVQRPIMSDLLLSSIFTAPTMVR
LLKGPWLRNPQYLATGILGAVVAMLLLNGLWPAYDDDFVIGGVTGIFGSWAGMALFDA
ILGIA"
gene 411. .2480
/gene="mlr1245"
CDS 411. .2480
/gene="mlr1245"
/codon_start=1
/transl_table=11
/product="probable electron transfer protein"
/protein_id="BAB48663.1"
/db_xref="GI:14022053"
/translation="MNSPANITDPLVLFMPSGKRGRFPVGTPVLDAARQLGVYVESVC
GGRATCGRCQIEVQEGNFAKHKIVSSNDHISPKGPKEERYERVRGLPERRRLSCSAQI
LGDLVIDVPQDTVINAQTIRKDADTRVIARDTAIRMCYVEIEEPDMHKPLGDLDRLKI
ALMKDWGFKNLEFDFYLLPQVQGILRKGNWTATAAIHKDADSDIARVIALWPGLKNEA
YGLACDIGSTTIAMHLVSLLSGRVAASSGTSNPQIRPGEDLMSRVSYVMMNPDGREGN
TVAVREAISSLVDKVCAEGNVQRNDILDSVFVGNPINHHLFLGIDPTELGGAPFALAV
SGAVRIKASDIGLKLNQGARLYMLPCIAGHVGADAAAVTLSEGPHRQDEMMLIVDVGT
NAEIVLGNRARVVAASSPTGPAFEGAEISGGQRAAPGAIERVRIDPDTLEPKYRVIGS
ELWSDEPGFLDSVQATGVTGICGSGIIEIVAEMYLAGIISEDGVVDGSLAARSPRVTA
NGRTFSYVLKEGEPKITITQTDVRAIQLAKAALYAGTKLLMEKQNTEHVDRIHFAGAF
GSFIDPKYAMVLGLIPDCDLDKVSAVGNAAGAGARMALLNRGYRREIEETVSQIEKIE
TALEPKFQEHFVYAMALPNKVDPFPKLSAAVKLPPRKTVSEDGIAGDATPRRRSREGH
AARRSRE"
gene 2529. .2852
/gene="mlr1246"
CDS 2529. .2852
/gene="mlr1246"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB48664.1"
/db_xref="GI:14022054"
/translation="MFPDRTDPETETLFSRNLSDTWQAHLPRIERRPEIILGRRSSRE
MKMSMFENLGRFGVTIRQNHARNKAIRALNSLPAHVQKDIGWPASPPSDPQAALASLL
LGTAR"
gene 3087. .4040
/gene="mlr1247"
CDS 3087. .4040
/gene="mlr1247"
/codon_start=1
/transl_table=11
/product="esterase"
/protein_id="BAB48665.1"
/db_xref="GI:14022055"

```
                     /translation="MPSLKSHPVSPVLRYSRKKAPSSPENLRRWIAAARKTEDHHPPA
                     SLHQRLHIQTRTVDGFPVYEIAPKAGEQRRILYLHGGAYVPEITPYHWHLIADMADRL
                     GYGITVPIYPIAPEHDFHAMPGMVGDVYRQMLDETEAEDIVFMGDSAGGNMAVVLTMM
                     AAEEGLPLPARHVLISPGLDMSLSNPKLFEAERNDPNLGIAGGLEAIRNYSAGIDRSD
                     WHISPIYGDVSVLPTTLLLTGSHDLLTPDNLIFAQKARDAGVEVEVVHEEGMFHAWPL
                     IDMPEARRARQHIVAPLGENRSPVRQDGKTRFQAPSAEAAE"
     gene            complement(4062..4961)
                     /gene="mll1248"
     CDS             complement(4062..4961)
                     /gene="mll1248"
                     /codon_start=1
                     /transl_table=11
                     /product="probable transcriptional regulator"
                     /protein_id="BAB48666.1"
                     /db_xref="GI:14022056"
                     /translation="MNAPLNHPLPLLDLDVLRTFVAIAETGSFTTAANAVFRTPSAVS
                     MQIKKLEDILGRSVFARDARSVSLTTDGEMLLGYARRLLSINREVVSKFIIPEIVGVV
                     RLGSPDDYGERVLPHVLKRFAQSHPSIAVDVTIDQSSNLRRRMDDRALDITLLTNSYK
                     TSALGAEVLLTEPIVWAGAKGGCAHLREPLPVSLWEEGCAWRAGALEALGREGRNYRV
                     AYMSAHTAGQRAAIMADLAVAPLPRSFLGNDMVELCPKDGMPDIGTYNLAMVVAPDAS
                     APVKAVADHIRATFEVFRETGKF"
     gene            5113..5379
                     /gene="msr1249"
     CDS             5113..5379
                     /gene="msr1249"
                     /note="unknown protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB48667.1"
                     /db_xref="GI:14022057"
                     /translation="MGLKHAACKTSQIKRRNVPGPAPKENEMSILSSLGRIATEFSAA
                     RARYQTERAIRSLPIELQKDIGWPEASDTKTGTRNGVGSWAGAK"
     gene            6136..7137
                     /gene="mlr1251"
     CDS             6136..7137
                     /gene="mlr1251"
                     /codon_start=1
                     /transl_table=11
                     /product="probable transcriptional regulator"
                     /protein_id="BAB48668.1"
                     /db_xref="GI:14022058"
                     /translation="MNAIKHPIKRSPVFFLVPDFTMIAFATALDPLRSANRMLGYEAY
                     RWRLASIDGKPVRASNGVECAVNTSLEEERKKMAGPDRPNMAIVCSGVNVERYQNKSA
                     PAWLRBEYNRGVAVGGLCTGAHILAAAGLLSNKRCAIHWENLPGFSEAFPKANVFADL
                     FEVDQHIYTCAGGTAALDMMLKLIGDDPDESLVNRVCEQVLTDRVRSPTDRQRLPLRA
                     RLGVQNSKVLTIIELMEGNLSEPLSLIEIADHVDLSRRQIERLFRTEMGRSPARYYLE
                     IRLDRARHLLIQSSMPVVEVAVACGFVSASHFSKCYRELYARSPQQERVDRKQLLAA"
     gene            complement(7153..7956)
                     /gene="mll1253"
     CDS             complement(7153..7956)
                     /gene="mll1253"
                     /codon_start=1
                     /transl_table=11
                     /product="ATP-binding protein of ABC transporter"
                     /protein_id="BAB48669.1"
                     /db_xref="GI:14022059"
                     /translation="MEPGVSKLLIEGVSRTFAGVRGGLPVKALMPIDLAVAANDFITI
                     LGPSGCGKSTLLRIVAGLEAPSEGRVLLDDKAVTRPGPDRGMVFQSYTLFPWLTVSEN
                     IAFGLRERGMAPKERDDIVASYVDLVGLKGFENHWPKQLSGGMQQRTAIARALANDPE
                     ILLLDEPFGALDNQTRGLMQELLLGIWERRKKTVLFVTHDIEEAIFMASRVIVMTARP
                     GSIKSDVAIDLPHPRHYTLKTSPEFSALKARLTEDIRVEAMRTAQAQPA"
     gene            complement(7941..8720)
                     /gene="mll1254"
     CDS             complement(7941..8720)
                     /gene="mll1254"
                     /codon_start=1
                     /transl_table=11
                     /product="permease protein of ABC transporter"
                     /protein_id="BAB48670.1"
                     /db_xref="GI:14022060"
                     /translation="MKMRPLHPVSPGLRTGLGISFFVLFIAFWANITLGGHVNRIFLA
                     DPLSMLRDGWRLLVEDRFWLDILITIWRVFGGFVLASVVAVPIGIAMGAWKPVEAPLE
                     PFVSPARYLPASAPIPLLILWAGIGELEKLLVIFVGSVFQIILIVAVKVGSTRRDLVE
                     AAYTLGSTNNGIVKRVIMPANAPEIAETLRLVLGWAWTYVIVAELIGSSSGIGYMIIN
                     SQSRLATGQIIFGIIVIGLIGLLSDFAFKAFNRWLFPWSLA"
     gene            complement(8802..9752)
                     /gene="mll1256"
     CDS             complement(8802..9752)
                     /gene="mll1256"
                     /note="hypothetical protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB48671.1"
                     /db_xref="GI:14022061"
                     /translation="MNRIAVFAATLTFAAGLSAPVMAATSVTIGISGWTGFAPLTLAK
                     QAGLFEKHGLDVTLKKVPQASRPLAIASGDLQCAATTVETWLVWNASGVTTKQIFQLD
                     KSYGADGIVVRNDIKTVADLKGKNVASSAPGTSPYFLLAWVLNKNGMSTKDVTVVNLE
                     PDAAAQAFLAGQNDAAVTYEPFISAVRDKADQGHILATTLDYPMVLDTVGCTPEFLKA
                     NPDAAKALADSYFEALDLIKKDPQKSYEIMGADVKQSAKEFEDSAKYLKWADKADNQQ
                     FFTKEFQDFSKTAGALLLQMGLIKEAPDVATLADTSAVAN"
     gene            complement(9963..11000)
                     /gene="mll1257"
     CDS             complement(9963..11000)
                     /gene="mll1257"
                     /note="unknown protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB48672.1"
                     /db_xref="GI:14022062"
                     /translation="MSSISARPIRSKSPVSATSRFMTRRTSSRAADRSGGPRDLLQKP
                     GFEPGFCLFKQVVFAYPRGMSALSRARRLLSGCAALALMLWLMPGSRADEPVDVELVL
                     AVDVSLSMSADELEIQRHGYAAALTHDNVLQAIADGAHGKIAVTYVEWAGTTWQRVIV
                     PWTVIANRADAERVVAQLSAQPPNSARRTSISGALEFGSDLFAESGYEGTKRVIDISG
                     DGPNNQGAPVNLTRDGVVRQGIVINGLPLMTRGGLSGAYDVNDLDRYYSDCVIGGPGA
                     FMIPVNDWTQPPEAIRRKLVLELAGPASPQWAAEEADHPPVVLTQDKPAADCQIGEKN
                     WRNRNWMLDSR"
     gene            complement(10906..13467)
                     /gene="mll1258"
     CDS             complement(10906..13467)
                     /gene="mll1258"
                     /codon_start=1

  Query Match             10.3%;  Score 163.6;  DB 1;  Length 329709;
  Best Local Similarity   48.7%;  Pred. No. 3.1e-29;
  Matches  608;  Conservative    0;  Mismatches  589;  Indels   51;  Gaps    4;

Qy        145 ccgcgtccacctcatcggtattggcggagccggaatgtctggcgttgcccgaatcctgct 204
               ||   || ||  ||||||| ||||||||    ||| |||    ||| | |||     | ||||
Db     249307 CCTTGTGCATTTCATCGGCATTGGCGGCATCGGCATGAGCGGCATCGCCGAGGTGCTGCA 249248

Qy        205 tgcccgcggtaagacagtcactggttccgatgccaaagattcccgcaccttgcttccact 264
                   | |||| | |  ||    || |||||          ||   |  ||   |||  |  ||
Db     249247 CAATCTCGGTTACAAGGTGCAAGGGTCCGACCAGGCCGACAGCGCCAATGTGCAGCGGCT 249188

Qy        265 ccgcgccgtgggagccaccatcgcagtgggacacgctgcggaaaccttgagctttccgg 324
                ||||  |  |||   |     |   || || |||   || || || || |     | |||
Db     249187 GCGCGACAAGGGCATCGAATGCTTCGTCGGCCACAAGGCCGAGAATCTCGGCGATGCCGA 249128

Qy        325 cgaacttcccaccgtcgtggtgacctctttgccgccattccgcaagacaacccggaact 384
                               |||||| |   |   | || ||   | |   ||| ||||||||
Db     249127 G---------------GTGGTGGTCGTCTCGACGGCGATCAAGAAGACCAATCCGGAACT 249083

Qy        385 tgttcgtgcacgtgaagaaggcattccggttattcgtcgctccgatctgttgggcgaatt 444
                      || || ||| |      | ||| |   | || ||| |||| || ||| |||   |
Db     249082 GAAGGCCGCGCGCGAAAAGCTGCTGCCGATCGTGCGCCGCGCCGAAATGCTGGCCGAGCT 249023

Qy        445 gctggaaggctccacccaggtcttgatcgcgggtacccacggtaagacctccaccacctc 504
               | ||        |      |||     |||||   || || || || ||||| | || |||||
Db     249022 GATGCGCTTTCGCCAGGCGGTGGCGATCGGCGGCACGCATGGCAAGACGACGACGACCTC 248963

Qy        505 tatgtctgtggtagctatgcaggcagcgggcatggatccaagctttgctatcggcggaca 564
                |||   | |         | |  || |  ||  | ||||| | | |    |  |||||
Db     248962 GATGGTCGCGACCTTGCTCGAAGCCGGCGGCTCGATCCGACCGTCATCAATGGCGGCAT 248903

Qy        565 gctcaacaaggctggcaccaatgcgcaccatggaactggtgaggtctttatcgctgaagc 624
                 ||||       |||||||||||||| |     ||      || |||     |  | |  || ||
```

```
Db 248902 CATCAATGCCTATGGCACCAATGCGCGCATGGGCGACGGCGAGTGGATGGTGGTCGAGGC 248843

Qy    625 agatgaatctgacgcatcgctgctgcgctacaagccaaatgttgcagtggtcaccaatgt 684
          || ||    ||||   |  | ||    |    | |  |  | || || ||||||||  |
Db 248842 CGACGAGAGCGACGGCACCTTCCTCAAATTGCCGGCCGACATCGCCGTCGTCACCAACAT 248783

Qy    685 ggaaccagaccacctggacttctttaaaaacccctgaagcctacttccaagtgttcgacga 744
          || || || || || ||| || |    |      ||      | | | ||||  | |
Db 248782 CGATCCCGAACATCTCGACCACTATGGCAGTTTCGACAAGGTGCGCGAGGCGTTCCGCCA 248723

Qy    745 tttcgcaggacgcatcaccccgaacggcaagctggttgtgtgcctgaacgatcctcacgc 804
          ||||  |    | |  |     |||||      ||  ||||     || ||||| | |
Db 248722 GTTCGTCGAGAACGTGCCGTTCTACGGCTTTGGCGTGATGTGCACCGACCATCCTGAAGT 248663

Qy    805 agcggagctgggggagaggtctgtccgcaagggtatcaagactgttggttacggtaccgc 864
          || |||||  |    |    |    |    | |||  ||    ||  |      |
Db 248662 GCAGGCGCTGGTCGGCCGCATCGAGGACCGCCGCGTCATCACCTACGGCGAGAACGCGCA 248603

Qy    865 tgacgccgtacaggcacaccctgaggttccagcgatggctaccatcgtggattcccaagt 924
          | ||  || |       |    |       |||  || |||          ||| || |
Db 248602 AGCCGATGTGCGCTTCACCAACCACCGCATGGCGGGCGCCACC---------TCCGAATT 248552

Qy    925 tgtcgcagaaggcacccgcgccaccatcaacatcgatggacaggaagtatctgtgattct 984
          | ||  |   |  |  |  ||  | |  || | |  |    |     |    ||   ||
Db 248551 CGACGTGGTGATCCGCGACCGCAAGACCGGCAGCCAGAGCACGATTTCAGGCCTGCGCCT 248492

Qy    985 tcaaatccctggtgatcacatggtactcaacggtgcagccgcccctgctggccggatacct 1044
          |  || || ||    |||| ||   |||||    | |||||   | .  | ||   |
Db 248491 GCCGATGCCCGGCCGCCACAATGTCTCCAACGCCACCGCCGCGATCGCCGTCGCCGCATGA 248432

Qy   1045 ggtgggtggggacgtcgacaagcttgttgaaggcttgtcggatttctccggcgtgcgacg 1104
          | | ||  |  || |||   | |     | ||| ||||   ||| |||||||     ||
Db 248431 GCTCGGCCTGTCCGGCGAGGCGATCAAGAAGGGCCTGTCGTCCTTCGCCGGCGTCAAGCG 248372

Qy   1105 ccgctttgagttccacggtgctatcgagggcggcaaatttaatggcgctgctatttatga 1164
          ||||||                | | |   |||    |  || ||||  | | | |  ||
Db 248371 CCGCTT--------------CACCCATACCGGATCCTGGAACGGCGTCGATGTGTTCGA 248327

Qy   1165 tgattacgcacaccacccaacggaagtaactgcagtgctcagcgctgcgcgcacccgggt 1224
          |||||||   ||||| ||    ||  |  | || ||||| |  || |||||||
Db 248326 CGATTACGGCCACCATCCGGTCGAGATCTCGGCGGTGCTGAAGGCGGCGCGCA------- 248272

Qy   1225 gaaggccgctggaaagggccgtgtcatcgtcgcgttccaaccacatttatactcacgcac 1284
          |||    |||||||| |||||| |     |||||| ||    | | | ||
Db 248273 -----GCGCCACCAAGGGCCGCGTCATCGCCATTGCCCAACCGCACCGCTTCACCCGGCT 248219

Qy   1285 catagaattccaaaaggagttcgcgggggcactgtcactggcagacgctgccgtggtgct 1344
          |   ||  |    || |||||| | |       |       || ||| | |   | ||
Db 248218 CCACGATCTGTTCAACGAGTTCTCCGCCTGCTTCAACGACGCCGACACGGTGATTGTCGC 248159

Qy   1345 tgagatttacggagcgcgcgaacaaccggtggatggcgtgtcctcgga 1392
          | | || |  ||  ||||| | ||| | || |||||   ||||||
Db 248158 GCCGGTCTATGCTGCCGGCGAAGACCCGATCGACGGCGTCAGCTCGGA 248111
```

Search completed: February 15, 2002, 08:59:52
Job time: 4502 sec